Complements of the Verb *Forget* in Late Modern English

Katja Rintala
Tampere University
School of Language, Translation and Literary Studies
English Philology
Pro Gradu Thesis
October 2014

Tampereen yliopisto
Englantilainen filologia
Kieli-, käännös- ja kirjallisuustieteiden yksikkö

RINTALA KATJA: Complements of the Verb *Forget* in Late Modern English

Pro Gradu –tutkielma, 74 sivua
Lokakuu 2014

Tämä korpuspohjainen pro gradu -tutkielma tarkastelee verbin *forget* komplementaatiota 1700-luvulta nykypäivään. Tavoitteena on selvittää mitä komplementteja *forget* valitsee ja miten komplementtien määrät jakautuvat. Mahdolliset muutokset komplementaatiossa ajan saatossa ovat tutkimuksen kannalta kiinnostavia. Myös verbin saamat merkitykset ja niiden yhteydet komplementteihin otetaan huomioon.

Tutkielma on jaettu kahteen osaan. Aluksi keskitytään tutkielman teoreettisen viitekehyksen esittelyyn, jonka jälkeen on empiirisen aineiston analyysin vuoro. Ensimmäisessä osassa tuodaan esille korpuslingvistiikkaan yleisesti liittyviä seikkoja ja käsitellään kompelementaation tutkimuksen kannalta olennaisia teorioita, kuten valenssiteoriaa, sekä verbin ja adjunktin välistä eroa. Tutkimuksen teoreettisena lähtökohtana voidaan pitää Rohdenburgin Great Complement Shiftiä, jonka mukaan englannin lausekekomplementaatiossa on tapahtunut merkittäviä muutoksia viimeisten vuosisatojen kuluessa. Lisäksi esitellään tutkielmassa käytetetyt aineistot ja tarkastellaan, miten verbiä on käsitelty valituissa kielioppiteoksissa ja sanakirjoissa. Empiirisessä osiossa vuorostaan esitellään ja analysoidaan korpusaineisto kronologisessa järjestyksessä, ja tehdään päätelmiä ja vertailuja verbin komplementaatiosta ja käytöstä analyysin perusteella.

Empiirinen aineisto on kerätty kahdesta elektronisesta korpuksesta. Historiallinen data on peräisin Corpus of Late Modern English Text –nimisestä korpuksesta, joka on jaettu kolmeen osioon ja sisältää tekstejä vuosilta 1710-1920. Vertailun vuoksi autenttisen datan lähteenä on myös British National Corpus, jonka sisältämät tekstit edustavat nykyaikaa. Molemmat tutkimuksen lähteenä toimivat korpukset keskittyvät vain brittienglantiin, ja tämä tutkielma onkin alueellisesti rajattu koskemaan ainoastaan brittienglantia.

Tutkimus paljasti muutoksia verbin *forget* komplementaatiossa. Nominaalilausekkeiden osuus aineistosta on pysynyt korkeana läpi tarkasteltavan ajanjakson. Great Complement Shift -teorian mukaisesti infinitiivikomplementtien osuus on vähentynyt, joskaan vastaavasti gerundien osuus ei ole noussut odotetusti. Sen sijaan esim. *that*-lausekkeiden määrä on kasvanut. Tutkimuksessa löydettiin myös yhteyksiä sanan eri merkityksien ja komplementtien välillä. Toisaalta todettiin myös, että Bolingerin periaate eri kielioppirakenteiden merkityseroista ei toteudu verbin *forget* kanssa, sillä samaa merkitystä käytettiin toistuvasti sekä infinitiivilausekkeiden että gerundien kanssa.

Avainsanat: forget, verbi, komplementaatio, korpus

Table of Contents

1 Introduction

The aim of this thesis is to examine the complementation of the verb *forget* and its inflectional forms *forgets, forgetting, forgot* and *forgotten* in Late Modern English. The distribution of complements over the chosen period of time as well as any changes detectable in the patterns used will be of special interest. As Mair (2002, 126) states, corpus linguistics facilitate the study of the ongoing change in English, although the results of this kind of a study are more likely to offer statistical information than a set firm rules. Corpus evidence can, however, help to provide empirical proof on a subject (ibid.). Furthermore, Mair attests that a corpus-based study should be conducted with respect to the theoretical linguistic tradition (2002, 109). The present study focuses on one particular verb, *forget*, i.e. it is a head-based project. Furthermore, as noted above, this is a corpus-based diachronic study and the scope of the study is regionally restricted to focus only on British English.

The study will be conducted first by examining earlier work done on the verb *forget* with the help of the *Oxford English Dictionary* (the OED) and other dictionaries along with selected grammar books, in addition to which some central theories concerning complementation and factors bearing on complement selection of verbs will be introduced. Secondly, the information offered by earlier work on the topic will be compared with authentic corpus data. The corpus data for this thesis is retrieved from the Corpus of Late Modern English Texts (hereafter CLMET) and the British National Corpus (the BNC). The CLMET covers the period from 1710 to 1920 whereas the BNC offers more recent data.

Chapter 2 discusses corpus studies in general, along with some problems associated with using corpus material as data in scientific research. Furthermore, the corpora used in this thesis will be introduced. Chapter 3 focuses on the theory of

complementation as well as the distinction between complements and adjuncts. In addition to that some factors bearing on the selection of complements will be introduced, such as the extraction principle. Chapter 4 examines previous work on *forget* with the help of two dictionaries and a selection of grammar books. Chapter 5 then discusses the primary data drawn from the two corpora that were chosen as sources for empirical data for this study. By examining the secondary sources and theories concerning the issue, and combining the theoretical knowledge with the authentic corpus data evidence, in the concluding chapter the present thesis will build an understanding of the senses and complement of the verb *forget* in Late Modern English period. In addition to presenting the frequencies and connections that are found between sense and complement, the possible changes taken place or in progress in the complementation system of *forget* during the recent centuries will also be of particular interest.

2 On Corpora

As this thesis is a corpus-based study and thus the primary data comes from corpora, I will begin with a chapter on corpora. Firstly I will introduce and discuss corpus studies in a general level. In addition, this chapter discusses some pitfalls involved in the use of corpus material and offers an introduction to the two corpora chosen as the sources of corpus data for this thesis.

2.1 Corpus studies in general

Thanks to the developing computer software and the spread of the internet and freely available corpora, corpus studies have become more and more popular since their first emergence in the early 1960's (Svartvik 1992, 7-11). The term *corpus linguistics* generally speaking refers to "the use of large collections of text available in machine-readable form" (ibid.). Leech (1968, 88) states that in the scientific study of a language we need to consider theory, description and data collection. *Theory* refers to the established form of language as "a phenomenon of the human mind and of society" while *description* refers to the established form of a given language with respect to the theory (ibid.). Consequently a description provides grammatical rules i.e. how the language should be used (ibid.). The third item, *data collection*, provides means to confirm a description and hence a theory through the collection of facts or observations (ibid.). Leech explains that the nature of linguistic data is somewhat controversial, and goes on to introduce three sources of data: corpora of actual utterances or texts, the elicited reactions of speakers of the language and the introspections of a native speaker of a language (ibid). Not being a native speaker of English, I will use corpora as a source for data in this thesis as the scope of elicited reactions and introspections is out of bounds for me. Svartvik further notes (1992, 9)

that corpora offer objective information (as opposed to introspection) which ensures the verifiability of the linguistic research. Svartvik reminds the reader that *linguistics* is indeed often regarded as the "scientific study of language", and introspection and elicitations of native speakers alone are not enough to study the linguistic competence and performance but instead it is vital that also real data (such as corpus data) is used (ibid.). Furthermore, for diachronic studies the use of corpus evidence is crucial, since there are no living informants available.

Svartvik claims that corpora provide the possibility of full accountability of linguistic features (1992, 9). However, as Leech notes (1968, 94), theory and practice do not always meet perfectly. Leech argues that no matter how large a corpus is it can never consist of all possible sentences in a language. However, he adds that it is important to remember that this does not make corpora less important as sources of empirical information as complete verifiability is not the goal in the testing of scientific theories. Moreover, Leech suggests that a corpus linguist studies "competence through the observation of performance" (ibid.). Mair continues that grammatical change can be detected in the "surface level changes" (2002, 108), i.e. it is worthwhile to study the frequencies of constructional variants (such as the complements of *forget*) and the shifts in the statistics. Furthermore, Mair states that the focus should be on documenting the gradual spread of a given construction and analyzing corpus evidence is a convenient way to get started (2002, 109).

Although corpora provide vast amount of empirical data on the language actually used, there are some factors that should be taken into consideration when conducting a corpus-based study. As Ball explains, when natural language is analyzed electronically, the performance is nowhere near perfect and the analyst should always be aware of the pitfalls involved (1994, 295). The possibility to use large on-line electronic corpora enables large masses of data to be analyzed in a relatively short

amount of time, but when analyzing corpus material, recall and precision can prove to be a problem. Salton, quoted by Ball (ibid.) has suggested that "precision is the proportion of retrieved material that is relevant, and recall is the proportion of relevant information that was retrieved". The problem with recall is that it is difficult to know how many relevant tokens were left out of the sample without manually examining the whole corpus, which is virtually impossible. The level of precision is more easily observable: irrelevant items are often recognized by the analyst and thus poor precision can be dealt with simply by omitting the irrelevant tokens from the sample (ibid.).

Furthermore, the raw frequency counts of features under investigation are not necessarily directly comparable. If the raw frequency of some linguistic construction is smaller in one text than in another, but the text in which the frequency was calculated is shorter than the other, using the raw number of instances of the given feature will not lead to accurate comparison. Instead, the frequencies should be normalized as Biber et al. (1998, 263) suggest. According to Biber et al. (ibid.), "the raw frequency count should be divided by the number of words in the text, and then multiplied by whatever basis is chosen for norming". Calculating the normalized frequencies enable exact comparisons regardless of the length of the text. As the corpora used for this study are also different in size as the following sections will reveal, I will use Biber et al.'s model of calculating the normed frequencies in the up-coming chapters discussing the primary data.

2.2 Corpus of Late Modern English texts

The Corpus of Late Modern English texts (later on the CLMET) is compiled by Hendrik de Smet on the basis of texts found freely on the internet using *the Project Gutenberg* and the *Oxford Text Archive* as mediums (de Smet 2005, 70). All of the

texts are essentially British English, which is suitable for this thesis because the study is restricted to British English only. The corpus consists of mostly imaginative literary texts, but there are also some formal or informative texts, e.g. Samuel Johnson's *Parliamentary Debates* (Vol. 1) (1740-41) and Walter Bagehott's *The English Constitution* (1867). The CLMET extends from 1710 to 1920 and it is divided into three sub-periods i.e. 1710-1780, 1780-1850 and 1850-1920 (ibid). All three parts are employed in this thesis to fully account for the development in the complementation of *forget* during the recent centuries.

The benefits of using the CLMET as a resource for data in this thesis is that it is considerably large, containing 9,818,326 words, therefore allowing deductions about the use of the different complements of the verb *forget*. By using all three parts of the corpus to cover the whole period from 1710 to 1920, some tendencies towards change in the use of the complements should become evident. De Smet has toiled to increase homogeneity of the texts within each of the three sub-corpora to facilitate detecting different historical trends (de Smet 2005, 70). Even though de Smet has knowingly focused on insuring variation among the texts, the corpus is biased towards male formal writers (2005, 78-79). As de Smet points out, this tendency may cause a situation where some changes in the language may not fully be detected in the texts, since the "linguistically self-conscious" authors often produce "the type of texts where one expects language change to be kept at a tight leash" (2005, 79). Furthermore, the predominance of male, highly educated, native British English speaking authors unfortunately results in the corpus being disproportioned in terms of gender, genre and register (2005, 78). However, granted that the present thesis will focus on British English only, the lack of regional variation is not critical. Moreover, the composition of the corpus is nevertheless consistent over all three parts, which ensures that the conclusions drawn from different parts of the

corpus are still mutually comparable (ibid, 79). As is mentioned above, the CLMET is restricted to written language and as Ball states, the results of a corpus study "can only be generalized to the extent that the corpus is a representative sample" (1994, 295). As a result, this present thesis will shed some light on the complementation of the verb *forget* in the written British English only and the analysis of the issue in spoken language will be disregarded.

2.3 The British National Corpus

Whereas the CLMET covers the period from 1710 onwards, the British National Corpus (the BNC) was chosen as the source for more recent corpus data. The information of this section is drawn from the BNC website accessed through http://www.natcorp.ox.ac.uk. The BNC contains roughly 100 million words from over 4,000 texts covering the late twentieth century. It was first released in 1995 and the latest revised edition, *BNC XML Edition*, is from 2007. The corpus was compiled by a consortium of major academic institutions, commercial publishing houses and public institutions. The BNC is a general corpus i.e. it is not restricted to any specific register or genre. It contains both spoken (10 % of the corpus) and written (90%) examples. The written texts represent both imaginative and informative texts. Texts for the corpus were selected so that the corpus would represent British English in its entirety and not just certain styles. The different text types are chosen to enable comparisons and contrasts between different types of texts. The selection was made in respect to domain, medium and time and these features were then further divided into classes, for instance 'domain' into 'imaginative' and 'informative'. 75% of the texts are from the informative domain and the remaining 25% represent the imaginative text type. When doing searches on the corpora, it is possible to choose only some of the domains and for example discard all spoken examples and concentrate on the written

language only. This feature is rather convenient as regards this thesis, since it is possible to choose such domains that best match the CLMET text types in order to ensure that the modern corpus data corresponds well with the historical corpus data.

The purpose of the corpus was to facilitate e.g. academic linguistic research and information retrieval. Furthermore, the corpus was designed to offer information on e.g. lexical, semantic, syntactic and morphological features of the language. As is the case with the CLMET, the BNC is also restricted to British English only. That suits the purposes of this thesis, because I chose to forget regional variation in the thesis and rather focus only on British English. Moreover, the focus is on the diachronic variation, and that is why the CLMET and the BNC complement each other so well.

3 On Complementation

As the title of this thesis suggests, verb complementation is the main theoretical framework of this thesis. This chapter introduces some central concepts regarding complementation in general. In addition to that, some complexity factors are discussed in more detail.

3.1 Complement vs. adjunct

Valency theory is a grammatical model focusing strictly on the verb. The main interest is in the verbal element of a clause and its relationship to its immediate subordinates (Somers 1987, 5). These elements are then further divided into complements, which are "expected to accompany a certain verb, or to complete its meaning" and adjuncts, which are in principle optional elements that "complete the meaning of the central predication as a whole" (Somers 1984, 508). The *valency* of a verb essentially refers to the amount of complements it takes (ibid.). In what follows a closer look will be taken at the distinction between complements and adjuncts.

As this thesis focus on the (post-verbal) complementation of the verb *forget*, let us begin with a definition of 'a complement'. Huang (1997, 75) suggests that complements are elements that "help complete the meaning of a sentence as required by a verb". Thus a sentence without a complement (e.g. *John saw*) is often regarded as "incomplete" and a complement (*John saw Bill*) completes the sentence (ibid.).

According to Huang (1997, 75), a verb can be followed by either a complement or an adjunct. The main distinction between complements and adjuncts is that a certain verb c-selects only certain complements whereas adjuncts can appear freely after any verb (ibid). Adjuncts serve to "provide additional information to

modify their heads, but are not required by the heads" (Huang 1997, 77). Huddleston (1984, 177) confirms that complements and adjuncts both follow matrix verbs but only complements are dependent of the head verb of the clause and only certain type of complements typically appear with certain verbs. He adds that only adjuncts can be omitted from a clause without changing the meaning or grammaticality of the matrix verb. Huddleston demonstrates the issue by an example sentence:

(1) *Unfortunately, my uncle was using an electric drill at that very moment*.

In (1) *My uncle* is the subject, *was using* the predicate and *an electric drill* the object i.e. the complement. If *use* is replaced by another verb, e.g. *arrive*, that selects different kinds of complements, the sentence becomes ungrammatical:

(2) **My uncle was arriving an electric drill*.

Adjuncts such as *at that very moment* or *unfortunately* can be added or omitted in both cases without the sentence becoming ungrammatical (Huddleston 1984, 178). As Huddleston and Pullum note, complements are more closely related to the verb than the freely occurring adjuncts (2002, 219). *Subcategorisation* is a term used to refer to the dependence between the verb and its complements. Different complementation patterns appear with different subcategories of verbs (e.g. 'monotransitive' or 'intransitive'). To put it in other words, different verbs license different complements and verbs are subcategorized according to the type of complementation they take. (ibid. 219-220). Huddleston and Pullum further note that typically verbs allow for multiple complementation patterns (220). There are also semantic restriction as to what kind of complements a verb licenses for. The following example sentences from Huddleston and Pullum (2002, 227) demonstrate how a basically grammatical sentence is odd when the selection restrictions are violated:

(3) a. *Kim enjoyed the concert*.
b. **The cheese enjoyed the cool breeze*.

In (3) b. the subject is –Animate while in (3) a. the subject is +Animate, and intuitively we see that the verb *enjoy* selects +Animate subjects rather than –Animate ones. The same applies for objects as well:

(4) a. *They frightened the cat*.
b. **They frightened the ironing-board*.

Clearly *frighten* selects +Animate objects rather than –Animate ones.

It is also worth noting that elements are always ascribed complements or adjuncts with respect to some verb (Somers 1987, 8-9). For example *in London* can be either a complement or an adjunct, depending on the matrix verb:

(5) a. James lives *in London*.

b. He looked for his friend *in London*.

Furthermore, the complement is sometimes obligatory in a sentence not only because its omission may lead to ungrammaticality, but also because its omission can in some cases change the meaning of the verb (Huddleston 1984, 179). Consider the following sentences:

(6) a.*He drives the minister mad*.

b.*He drives the minister*.

Both are grammatical enough but the sense of the verb *drive* changes (ibid.).

To summarize, the division of arguments into complements and adjuncts is of great importance in valency theory, and adjuncts differ from complements in that they occur relatively freely and the matrix verb does not determine their form (Herbst et al. 2004, xxiv). However, there are some known tendencies and factors concerning the choice of the complement, and these will be discussed in the next section.

3.2 Complexity factors

There are a number of complexity factors bearing on the complement selection of verbs. As this thesis focuses on the post-verb complementation of the verb *forget*, it is crucial to examine these extra-semantic factors such as extractions, insertions and horror aequi and later on consider what effect these factors have on the data drawn from the corpora. According to Vosberg (2009, 212-213) the competition between infinitives and gerunds is a widely known phenomenon in English linguistics and there has been a substantial tendency to replace infinitives by gerunds in the complement function. Rohdenburg has also contributed to the study of the change and refers to it as the Great Complement Shift (2006, 143). As regards the present thesis, claims about the gerund form becoming established on the expense of infinitival forms and *that*-clauses will be of particular interest when analyzing the corpus data later on. All of the extra-semantic factors to be discussed have a role in either hastening or stalling the change.

Furthermore, Bolinger's generalization states that "a difference in syntactic form always spells a difference in meaning" (1968, 127). According to Bolinger, infinitival forms are used in reference to the hypothetical, potential or imaginative while *–ing*-forms refer to something real, concrete and reificated (1968, 123-124). Bolinger argues that the *–ing*-form is probably advancing at the expense of the infinitive (1968, 125). In respect of the generalization the possible semantic differences among the tokens to be analyzed shall be accounted for as well.

3.2.1 Complexity principle

Rohdenburg's (1996, 151) complexity principle (or "transparency principle") declaims that "in the case of more or less explicit grammatical options the more

explicit one(s) will tend to be favored in cognitively more complex environments". That is, when there is a complexity factor present in a sentence, the *to*-infinitive is expected to be favored on the expense of the –*ing*-form. Using the complexity principle presumes that we can distinguish between the more and less explicit options, and accordingly Rohdenburg suggests that "the more explicit variant is generally represented by the bulkier element or construction" (1996, 152).

The cognitively more complex environments include among others finite complements (as opposed to infinitival complements) as well as environments of extractions, insertions and horror aequi which will all be discussed in more detail in the following sections (1996, 176).

3.2.2 Extractions

Vosberg's extraction principle is a central contribution to the study of change in English complementation system:

> In the case of infinitival or gerundial complement options, the infinitive will tend to be favoured in environments where a complement of the subordinate clause is extracted (by topicalization, relativization, comparativization, or interrogation etc.) from its original position and crosses clause boundaries. (Vosberg 2003a, 308).

Vosberg gives a list of different types of extraction structures (2003b, 201-202):

Topicalization	even her acquintance with the Belfield's$_i$
	she remembered [not ever mentioning t_i]
Relativization	it is the worthy Spencer$_i$ whom$_i$ I'm sure you remember [to have
	often heard [me mention t_i in the relation of my private misfortunes]]
Clefting	It was the bangle$_i$ that he remembered [having seen t_i on Francie's wrist]
Comparativization	Twas her Charming Face and modest Look, that represented to him a thousand
	more Beauties and taking Graces$_i$, than he remembered ever [to have seen t_i
	in his Unconstant and Faithless Mistress]
Interrogation	Now, how many$_i$ do you remember [to have heard named t_i]?

Table 1. Types of Extractions.

In the example sentences in table 1. the *gaps* (the original position of the object) are marked t_i and the *fillers* (the extracted element) are marked $_i$.

As regards the present thesis, in the corpus data later on it is of interest to study all cases of extraction involved. According to the extraction principle, in the tokens where there are traces of extraction, the infinitive form should appear more frequently than the –*ing*-form. Correspondingly, extractions are not to be expected among tokens with gerund complements.

3.2.3 Insertions

Rohdenburg claims that in respect of the complexity principle, 'structural discontinuity' in a sentence should attract the more explicit finite complement (2006, 148). By structural discontinuity Rohdenburg refers to what will be considered insertions when analyzing the data, i.e. any intervening element between the subject and the matrix verb:

(7) He hesitated for a very long time about whether he should do it / whether to do it.

In other words, when insertion is involved the *to*-infinitive should be more likely to appear than the gerundial complement.

3.2.4 Horror aequi

Another extra-semantic factor affecting the distribution of complement variants is the horror aequi principle. According to Rohdenburg (2003, 236) the horror aequi principle is a "widespread (and presumably universal) tendency to avoid the use of formally (near-) identical and (near-) adjacent (non-coordinate) grammatical elements or structures." In other words, two –*ing* forms are not likely to appear one after another and vice versa, a *to*- infinitive is not expected to follow a *to*- infinitive etc.

According to Rohdenburg (2003, 235), the horror aequi principle can result in some linguistic features to resist change longer if the environment is unfavorable, or on the other hand, some others to advance more rapidly if the environment is favorable. Indeed, it is a question of tendency rather than a firm rule.

4 Earlier work on *forget*

In this chapter I will investigate what has been said on the verb *forget*, its complements and usage previously in dictionaries and by grammarians.

4.1 *Forget* in dictionaries

First I will take a look at *The Oxford English Dictionary* (hereafter the OED). Since this is a diachronic study, the senses a word has taken in the recent centuries are also relevant. However, there were some sub-senses labeled 'rare' or obsolete or there were no recorded instances after 1710, and these were left outside the scope of this study. An example of such a sense would be e.g. OED sense 2c. 'with complementary adj. or adv. *Obs*': "Hys feris all hes hym forȝet allane" (1513 DOUGLAS). Even the example sentence given in the OED is from the 16th century and the sense is labelled obsolete, so the sense is altogether disregarded from the present study as it is not expected to occur among the corpus data, since all the data from the corpora used in this thesis is drawn from texts written and published only in the 18th century or later.

The OED lists five different senses for the verb *forget*, and a total of ten sub-senses under them. Even with the irrelevant senses disregarded, the senses offered in the OED are very elaborate and manifold. Table 1 below introduces all the relevant senses and sub-senses found as well as example sentences given by the OED. Furthermore, the complement or complements offered by the OED are presented in the right-hand column of the table.

Sense	Example	Complement
1. *trans.* To lose remembrance of; to cease to retain in one's memory. Often with clause as obj.	1845 S. AUSTIN Frederic . . Did not forget his numerous wrongs and affronts.	NP
	1874 GREEN Men forgot how to fight for their country when they forgot how to govern it.	*wh* -clause
	1757 WESLEY Have you forgot that every man is now born in as good a state as Adam was made at first?	*that* -clause
b. To fail to recall to mind; not to recollect	1787 'GAMBADO' He says much the same of rabbits and onions, but I forgot (? *read* forget) how he brings that to bear.	*wh* -clause
d. *absol* . (or *intr.*) Also, *to forget about* : not to recall the facts concerning; not to remember to take action in the matter of.	1897 N. GOULD He had forgotten about that, it was such a long time ago.	*about* + NP
2. To omit or neglect through inadvertence. Chiefly with *infinitive* as obj.	1842 TENNYSON The steer forgot to graze.	*to* -infinitive
b. To omit to take, leave behind inadvertently.	1768 GOLDSM. I had almost forgot the wedding ring!	NP
d. To omit to mention, leave unnoticed, pass over inadvertently.	1548 HALL I may not forget how the Frenche men . . .	*wh* -clause
	1881 FREEMAN Not forgetting a gate which has been made out in the long walls.	NP
3. To cease or omit to think of, let slip out of the mind, leave out of sight, take no note of.	1888 MISS BRADDON Are you forgetting luncheon?	NP
4. In stronger sense: To neglect wilfully, take no thought of, disregard, overlook, slight.	1703 BURKITT Men wallow in wealth, and forget God.	NP
5. *to forget oneself* . b. To lose remembrance of one's own own station, position or character; to lose sight of the reqirements of dignity, propriety or decorum; to behave unbecomingly.	1794 NELSON These Agents forget themselves very much.	reflexive
d. To lose consciousness.	1717 POPE *Mod* . I was nearly asleep, I had just forgotten myself.	reflexive

Table 2. *Forget* in the OED.

As can be detected from the table, *forget* can appear with several different complements. The most common one seems to be NP, which is used in accordance with many of the sub-senses. In addition, *forget* can be used with *about* + NP, *wh*-clause, *that*-clause, *to*-infinitive or reflexive complements. As the senses seem so similar to each other at first glance, a brief analysis of them is convenient. Sense 1

seems to apply in cases where one does not have remembrance of something. Sense 2 involves an aspect of inadvertence; one accidentally forgets something or omits something by mistake. Sense 3 is somewhat similar to sense 1, but opposite to sense 1, there is the idea of retaining something in one's memory and only momentarily forgetting it. Furthermore, sense 4 differs from the other senses in that it involves a conscious decision to forget something; to choose to overlook or disregard something. Then again, sense 5 is idiomatic and refers to inappropriate behavior.

I decided to include the *Collins Cobuild English Language Dictionary* (hereafter referred to simply as Collins/Cobuild) in this study for further reference. The following Table 2 was compiled according to the same standards as Table 1 above, so that the differences between the two dictionaries are easily detectable. I decided to include the example sentences in the table, since their role in the dictionary is somewhat smaller than in the case of the OED. Some of the senses share the definition and complements offered by the Collins/Cobuild and differ only in the level of abstractness, so they are treated as one here.

The Collins/Cobuild sense and example sentences	Complements offered
1. 'Not to remember' I never forget a face or a name . . . . . .that she had forgotten how to ride a bicycle	NP, *that*-clause, gerund, *wh*-clause
2. 'To overlook' He often forgot what he was supposed to be doing I forgot all about him for several years . . .	NP, *that*-clause, *wh*-clause, *about* + NP
3. 'To omit', 'to neglect' I meant to see her . . . but I forgot all about it. He forgot my birthday again this year. Sorry to disturb you - I forgot my key.	NP, *that*-clause, *about* + NP, infinitive
4. 'To reject' . . . I think you ought to forget her. . . . We can forget any idea of competing with . . .	NP, *that*-clause, *about* + NP, infinitive
5. 'To misbehave' "Oh darling!", cried Judy, forgetting herself.	Reflexive

Table 3. The verb *forget* in Collins/Cobuild dictionary.

According to Collins/Cobuild, *forget* can take the pattern V + O (i.e. 'object') / report clause, which in simpler terms means a NP complement or a *that*-clause. In addition to that, it can also take an *–ing*-clause, a *to*-infinitive, a reflexive pronoun or an *about* + NP complement. It is worth noting, that the OED does not refer to the gerund as a possible complement of *forget* whereas it is mentioned in Collins/Cobuild. In the example sentence Collins/Cobuild gives in reference to the previously mentioned variant, there is an intensifying element *all* inserted, which provokes thoughts on the frequency of such insertions, and therefore it will be of interest in processing the corpus data to observe how frequently *all* tends to be inserted in such environment.

The various senses given by the two dictionaries are somewhat overlapping and therefore I took the privilege to combine and simplify the meanings to suit the purposes of this thesis better. The OED senses 1 b and d together form the simplified sense 1, 'not to remember or fail to recollect'. Sub-senses of the sense 2 are combined to form the simplified sense 2 'to omit or neglect through inadvertence'. Sense 3 is simply 'cease or omit to think of' whereas sense 4 is 'disregard or overlook'. Sense 5 is simply 'to forget oneself'. Below is a table of the simplified senses based on the OED entry for s.v. *forget*, and as can be seen, it rather resembles the table of Collins/Cobuild senses even though the complements offered are somewhat different.

The OED senses simplified	Complements offered in the OED
1. 'not to remember or fail to recollect'	NP, *wh*-clause, *that*-clause, *about* + NP
2. 'to omit or neglect through inadvertence'	NP, infinitive, *wh*- clause
3. 'to cease or omit to think of'	NP
4. 'disregard or overlook'	NP
5. 'to forget oneself'	reflexive

Table 4. Simplified senses

To summarize, the two dictionaries give a total of seven complements for the verb *forget*; NP, *about* NP, infinitive, gerund, *that*-clause, *wh*-clause and the reflexive pronoun. In order to discover any other possible complements and to shed more light on the issue, the next section will be examining grammar books. Finally, after discussing all secondary sources, a comprehensive table of all the possible variants offered by the dictionaries and the grammar books will be compiled.

4.2 *Forget* in grammars

In this section a selection of grammars will be consulted to see how the verb *forget* and its complements should behave according to them.

According to Poutsma (1904, 618), *forget* requires the construction *to* + infinitive:

(1) I forgot <u>to answer</u> a question which you asked me

(2) The steer forgot <u>to graze</u>.

Poutsma does not mention any other possible patterns. The other grammarians, however, introduce several additional possibilities.

Leech and Svartvik state that *forget* is a verb that takes one object (i.e. *forget* is a transitive verb that selects only one complement at a time) and thus appears in SVO patterns. They argue that the object of *forget* is often *to* + infinitive, but it can also be followed by an *–ing* form, or a *wh*-clause (2002, 406-408). The following example sentences are given by Leech and Svartvik to illustrate the types of complements they offer:

(3) We agreed <u>to stay</u> overnight

(4) We ought to avoid <u>wasting money</u> like this

(5) She forgot <u>where to look</u>.

Quirk et al also note only these three patterns. Furthermore, they claim that the complement can be a "subjectless *ing*-participle clause" (1985, 1189), where the subjects of the non-finite verb and the preceding verb are usually the same, as in:

(6) I love listening to music.

However, the subject of the participle is not necessarily determined by the subject of the preceding verb, but can have indefinite meaning (ibid, 1190).

Quirk et al also state that with *forget* the –*ing*-clause is rare without preceding *about* (ibid, 1193).

(7) I forgot (about) going to the bank.

Biber et al also identify the pattern *about* + -*ing*-clause (1999, 742), according to which the pattern *forget* + *about* + -*ing*-clause should be more common than *forget* + -*ing*-clause. That may not be the case, however, because neither Leech and Svartvik (2002, 406) nor Huddleston and Pullum (2002, 1232) mention *about* at all in connection with the gerund. It will be interesting to see what the distribution of these patterns is in the corpus material.

Biber et al (1999, 661-663) as well as Huddleston and Pullum (2002, 1008) note that in addition to the patterns already mentioned above, *forget* can also take a *that*-clause complement.

Table 3 below will list the complements mentioned by each of the secondary sources more clearly:

	Poutsma	Leech & Svartvik	Quirk & al.	Biber & al.	Huddleston & Pullum	OED	Collins/ Cobuild
To + infinitive	+	+	+	+	+	+	+
Gerund		+	+	+	+		+
About +gerund			+	+			
About + NP						+	+
Wh- clause		+	+	+	+	+	+
That-clause				+	+	+	+
NP						+	+
Reflexive						+	+

Table 5. Complements discussed in secondary sources.

5 *Forget* in the Corpus of Late Modern Texts

In this chapter I discuss authentic corpus data collected from the Corpus of Late Modern English Texts (the CLMET) and the British National Corpus (the BNC) will be examined in chapter 6. The CLMET is divided into three sub-corpora, which are all taken into account in this study. There is also an extended version of the CLMET available, but as already stated above, the original version was used instead. It was sufficient for the purposes of this study, since *forget* is such a common verb that there was no shortage of occurrences for *forget* even in the original version. In fact, there were altogether 2 297 hits in the whole CLMET (all three parts) for *forget* and its inflectional forms, so I decided to study only a sample the size of one third of the total amount. That left me with a total of 765 hits, of which 434 will be under examination in this thesis. Irrelevant tokens, such as cases where *forget* is used as an adjective or has a zero complement, will be discarded from further study. Furthermore, passive constructions will be analyzed as they were in the active voice. I will discuss each three parts of the CLMET separately in sections 5.1, 5.2 and 5.3 and their subsections. After that, in section 5.4, I will present a brief summary of the findings and frequencies in each of the sub-corpora. After completing the analysis of the historical corpus I will turn to the BNC for more recent data. The results are then compared to each other and findings related to the secondary sources.

5.1 CLMET part 1

The first part of the CLMET contains 2,096,405 words from 24 different texts written in 1710-1780. To ensure sufficient recall, the search string included all forms of the verb i.e. *forget, forgets, forgot, forgotten* and *forgetting*. The query yielded a total of 336 tokens. In order to keep the sample manageable only every third token was

selected, so 112 tokens were left for further investigation. The level of precision was satisfactory as only five of these proved to be irrelevant; in three of the tokens the past participle forms of *forget* were used as adjectives and two had zero complements:

(1) ... hout lovers, old without a friend; A fop their passion, but their prize a sot; Alive, ridiculous; and dead, **forgot**! Ah! friend! to dazzle let the vain design; To raise the thought and touch the heart be thine! (Pope 1773-1774, *An Essay on Man*)

(2) "Not against thy father; indeed, I dare not," said Theodore. "Excuse me, Lady; I had **forgotten**. But could I gaze on thee, and remember thou art sprung from the tyrant Manfred! (Walpole 1764, *The Castle of Otranto*)

Since the aim is to study the complements of *forget*, these will not be analyzed further. Instead, I shall discuss the remaining 107 tokens and the complementation patterns found in them in more detail. Table 4 below illustrates the distribution of the different patterns found in the sample. Normalized frequencies are calculated as per one million words i.e. the raw frequency of tokens is first divided by the total count of words in the texts under study and then multiplied by one million (see 2.1). However, since I only use a sample size of one third of the whole corpus in the present study, it must be kept in mind that the number of words in CLMET part 1, which is 2,096,405 in total, must also be divided by three to make the calculations more accurate. Thus the number of tokens in the sample are divided by 2,096,405 : 3 = 698,801 and ´then multiplied by one million. To make the matter more simple, I use the pattern X : Y x 1,000,000 in which X= raw number of tokens in the sample and Y= number of words in the sample and one million is used as the basis of norming. This way the frequencies in different parts of the corpus are comparable to each other. The same method will be used throughout the thesis and will not be accounted for in such detail further on. Furthermore, the percentage shows the ratio of a certain type of complement to the total count of tokens in the current sample. The percentage is accordingly calculated by dividing the number of tokens of a certain type of

complement by the total number of the tokens. Again, this pattern will also be employed later on in the thesis, but will not be explained in detail after this. The normed frequency of the total number of the complements in the bottom row in the table refers to the overall prevalence of *forget* in the corpus as compared to all other words in the sample. Accordingly, Table 6. below shows all the statistics, and the following subsections will then discuss all of the different type of complements in more detail.

Type of complement	**Number of tokens**	**Normed frequency**	**Percentage**
Noun phrase	75	107,33	70
To + infinitive	20	28,62	18,7
That- clause	8	11,45	7,5
Wh- clause	4	5,72	3,7
Total	107	153,12	100

Table 6. The complement patterns of *forget* and their frequencies in the 1st part of the CLMET.

5.1.1 *Forget* + noun phrase

Although none of the grammarians discussed the pattern *forget* + NP, the non-sentential NP complement was by far the most common pattern in the present sample with 75 instances out of 107.

One interesting notion is that the frequency of negations seemed rather high among the *forget* + NP structure. Out of the 75 NP complements in the sample, 19 had a negative construction:

> (3) I have not **forgot** your magazines, but will send them and these pamphlets together. Adieu! I am at the end of my tell. (Walpole 1735-48, *Letters*)

> (4) But, though he gave no credit to my information, I had sufficient reason to think he did not **forget** it, by the resentment which he soon discovered to both the persons whom I had named as my informers. (Fielding 1751, *Amelia*)

According to Biber et al (1999, 159), the negative construction is indeed common with "mental verbs", such as *forget*.

After identifying all the tokens with NP complements, I analysed which sense of *forget* each token represents. The distinction between different senses was not always clear-cut. In some cases more context than just the short example sentence given in the corpus material was needed in order to determine the appropriate sense. Some of the example sentences are rather short while others are considerably long, and the length of the sentence depends on the amount of context needed to determine the sense. The count for each sense of the NP complements can be seen in Table 7. below:

OED sense	Frequency
1. 'not to remember or fail to recollect'	37
2. 'to omit or neglect through inadvertence'	17
3. 'to cease or omit to think of'	15
4. 'to disregard or overlook'	6
5. 'to forget oneself'	--
Total	75

Table 7. The senses of the NP complements in CLMET part 1.

According to the dictionaries NP complement is used in connection to various different senses, and the distribution of senses in Table 7. seems to conform to that as NP complement clauses were used widely in almost all senses of *forget*. As the table suggests, the most common sense is the OED sense 1 'not to remember or fail to recollect':

(5) 'I shall never **forget** the sneer with which she spoke those last words." (Fielding 1751, *Amelia*)

(6) was not Democritus, who laughed ten times more than I –town-clerk of Abdera? and was not (I **forget** his name) who had more discretion than us both, town—clerk of Ephesus? (Sterne 1759-67, *The Life and Opinions of Tristram Shandy*)

Sense 2 was the second most common variant in the sample. The sense of inadvertence was crucial when determining whether a token belongs to sense 2 or not, and it also includes all instances of leaving something behind unintentionally:

(7) I had almost **forgotten** one thing, which I would recommend as an object for your curiosity and information . . . (Chesterfield 1746-1771, *Letters to his Son*)

(8) I have not **forgot** your magazines, but will send them and these pamphlets together. (Walpole 1735-48, *Letters*)

Sense 3 'to cease or omit to think of' was almost as common:

(9) He smiled at the narrative of my travels, and was glad to **forget** the constellations and descend for a moment into the lower world. (Johnson 1759, Rasselas, *Prince of Abyssinia*)

There were no tokens with NP complement that would fall under sense 5 'to forget oneself'.

As the number of NP complements is so high compared to all other type of complements, the semantic qualities of the subject and complement of the clause were investigated further to learn more about the NP complement clauses. The nature of the subject was considered with regard to such features as +/– Human whereas the NP complement was deemed +/– Abstract. The subject was +Human in 72 of the tokens while only three had –Human subjects. The three –Human referred to war policies, faculties of the mind and imagination:

(10) It is obvious, that the imagination can never totally **forget** the points of space and time . . . (Hume 1739-40, *A Treatise of Human Nature*)

Furthermore, the NP complement represented +Abstract quality in 60 instances:

(11) . . . but this be assured of likewise, that whilst I live I shall never **forget** the kindness of the offer. (Fielding 1751, *Amelia*)

As a result, only 14 NP complements were –Abstract in nature. In ten of the cases with –Abstract NP complement, the NP was a personal pronoun or otherwise referred to human beings, as seen in example (12), and the four remaining ones included concrete things such as a palace, magazines, seeds and gout, as is seen in example sentence (13):

(12) .. . . but this is the way that all great men **forget** their friends . . . (Goldsmith 1766, *The Vicar of Wakefield*)

(13) Sir Thomas, waked by this tremendous exclamation, started up, and **forgetting** his gout, followed the lieutenant's example by a kind of instinctive impulse. (Smollett 1771, *The Expedition of Humphrey Clinker*)

As for extra-semantic features, there were four instances where a process involving extraction was evident. In determining which type of extraction is present in each case, the table given by Vosberg (2003b, 201-202) (cf. 3.2.2) was used as a basis of the present analysis. Accordingly, all four instances were cases of relativization:

(14) Now for your own commissions$_i$, which$_i$ you seem to have **forgotten**. (Chesterfield 1746-71, *Letters to his Son*)

(15) But having once acquired this notion of causation from the memory, we can extend the same chain of causes, and consequently the identity of car persons beyond our memory, and can comprehend times, and circumstances, and actions$_i$, which$_i$ we have entirely **forgot**, but suppose in general to have existed. (Hume 1739-40, *A Treatise of Human Nature*)

Insertion was present in two NP complements, the other involving a subordinate clause and the other an adjunct:

(16) I **forgot**, *after a time*, my disappointment . . . (Johnson 1759, *Rasselas, Prince of Abyssinia*)

5.1.2 *Forget* + *to*-infinitive

In addition to the non-sentential NP complements, forget also selects sentential complements the infinitival being the most common of them. According to Poutsma (1904, 618), this is the only possible complement of *forget*, and since it is the only pattern mentioned in all the works cited, I expected the pattern to be particularly common. However, although the *to* + infinitive was the second most common complement found in the corpus data, there were only a total of 20 tokens. Both dictionaries mention the *to*-infinitive complement only in connection to sense 2 'to omit, to neglect (inadvertently)', which seems to be consistent with the corpus material as all of the tokens represented the OED sense 2. As was mentioned earlier in section 3.2, infinitival forms typically refer to something hypothetical, potential or imaginative while *–ing*-forms are used in reference to something real, concrete and reificated (Bolinger 1968, 123-124). The unintentionality of the forgetting in sentences representing the OED sense 2 could be thought to decrease the level of concreteness of forgetting and thus Bolinger's claim would be confirmed.

What is worth noting here is that the complement in eleven out of the 20 instances was a verb of communication (*to tell, to say, to mention*). *Forget + to tell + NP* appeared seven times whereas *utter*, *speak*, *mention* and *say* occurred only once:

(17) I quite **forgot** to tell her that I intended to prepare at the bottom of the garden. (Goldsmith 1773, *She Stoops to Conquer*)

(18) ... s of Lampridius, he ought not to have followed the unjust severity of Herodian, and, above all, not to have **forgotten** to say that the virtuous Alexander Severus had insured to the Jews the preservation of their privileges, and ... (Gibbon 1773, *The Decline and Fall of the Roman Empire* Vol. 1)

The remaining nine tokens seemed to be rather random and consequently there was no particular pattern to be found. There was no evidence of extractions or insertion

present in the material. Moreover, there were no violations against the horror aequi principle.

It should be noted here, that there were no gerunds to be found within the present CLMET sample. According to Rohdenburg's Great Complement Shift theory (2006, 143), gerunds should be on the increase at the expense of the infinitival and *that*-clause complements, which is not in accordance with the findings so far as there are no gerunds and the most common sentential complements of *forget* are in fact infinitives and *that*-clauses, which will be discussed in more detail next. Therefore it will be of interest to see what the situation is like in the later parts of the CLMET and in the BNC data.

5.1.3 *Forget* + *wh*-clause / *that*-clause

There were four tokens with a *wh*-clause complement and eight with a *that*-clause complement. All grammarians, except Poutsma, note that *forget* can take a *wh*-clause complement. Quirk et al. even mention that the verbs that take *that*-clause complements usually also take *wh*-clause complements (1985, 1184), and with *forget* this seems to be the case.

As there were only four instances of the *wh*-clause, it is difficult to make extensive conclusions of the nature of *wh*-clauses chosen by *forget* based only on this data alone, as there were hardly any common nominators discernible among these example sentences. However, the results are nevertheless far from irrelevant as this serves as a good starting point and a basis of comparison for the latter parts of the CLMET as well as for the BNC data later on in the thesis.

In two of the *that*-clauses *that* was omitted, and in both the subjects of the lower and higher clause were the same:

(19) "Truce to your sermon," said Manfred; "you **forget** you are no longer Friar Jerome, but the Count of Falconara. (Walpole 1764, *The Castle of Otranto*)

Furthermore, there was insertion in one of the tokens with *that*-clause complement:

(20) “Nor have I **forgotten**, sir,” *said Theodore*, “that the charity of his daughter delivered me from his power. (Walpole 1764, *The Castle of Otranto*)

As for the senses that the tokens with *wh*-clause or *that*-clause complements, according to the dictionaries sense 1 is the only option given for *that*-clauses by the OED whereas there are multiple choices according to Collins/Cobuild. *Wh*-clauses are suggested to be common with senses 1 and 2. In fact, sense 1 proved to be the most common variant as almost all tokens had sense 1:

(21) When in the course of toying he dropped a purse into her bosom, she seemed to **forget** how the night wore, and, with the approbation of her charge, assented to his proposal . . . (Smollett 1751, *The Adventures of Peregrine Pickle*)

(22) But Jones unluckily **forgot**, that though the hand of Norherthon were tied, his legs were at liberty; . . . (Fielding 1749, *The History of Tom Jones, a Foundling*)

In addition to sense 1, even though neither of the dictionaries considers it possible, there was one *wh*-clause complement and one *that*-clause complement with sense 3:

(23) It is their excellences which have taught you their defects. I would wish you to **forget** where you are, and who it is that speaks to you. (Reynolds 1769-76, *Seven Discourses on Art*)

(24) A lame youth, whom Apollo had recompensed with a pipe, and to which he had added a tabourin of his own accord, ran sweetly over the prelude, as he sat upon the bank—Tie me up this tress instantly, said Nannette, putting a piece of string into my hand—It taught me to **forget** I was a stranger—The whole knot fell down—We had been seven years acquainted. (Sterne 1759-67, The Life an Opinions of Tristram Shandy)

The distribution of senses can be seen in tables 8 and 9:

OED sense	Frequency
1. 'not to remember or fail to recollect'	7
3. 'to cease or omit to think of'	1

Table 8. The senses of *that*-clause complements in CLMET part 1.

OED sense	Frequency
1. 'not to remember or fail to recollect'	3
3. 'to cease or omit to think of'	1

Table 9. The senses of *wh*-clause complements in CLMET part 1.

What is interesting, the token in example (24) with sense 3 also had *that* omitted, which raises questions regarding the possible connections with omission of *that* and different senses. Furthermore, according to Bolinger's generalization "a difference in syntactic form always spells a difference in meaning" (1968, 127), and at least in this case it seems to hold true. However, as there was only one token in the first part of the CLMET, no conclusions can be drawn merely based on this, but the issue certainly is worth further study later on in the thesis. Moreover, as Rohdenburg's complexity principle (1996, 151) suggests, the more explicit grammatical variant will tend to be favored in cognitively more complex environments. That is, when there is a complexity factor such as insertion present, *that* is expected to be present rather than omitted. These findings certainly conform to that, but again, further study is needed.

5.2 CLMET part 2

The second part of CLEMT contains 3,739,657 words from 40 different texts written in 1780-1850. The query for *forget*, *forgets*, *forgot*, *forgotten* and *forgetting* yielded a total of 967 hits, of which I took a sample containing one third of the hits. That left me with 323 instances. Further investigation proved that 36 occurrences were irrelevant for this study; *forget* (or rather its inflected forms *forgot* or *forgotten*) was

used as an adjective or had a zero complement in addition to which a token with a misspelling of the noun ‘forgetfulness’ was found:

> (25) Wild visions of enlistment, of drinking himself into **forget** fulness [sic], of becoming desperate in some way or another, entered his mind; . . . (Gaskell 1848, *Mary Barton*).

After excluding the irrelevant tokens there were 286 tokens to be taken under further investigation. The following table (Table 5) shows the distribution of complements:

Type of complement	**Number of tokens**	**Normed frequency**	**Percentage**
Noun phrase	209	167,66	73,1
That –clause	40	32,09	14,0
To –infinitive	20	16,04	7,0
Wh –clause	8	6,42	2,8
Reflexive	4	3,21	1,4
About + NP	3	2,41	1,0
Gerund	2	1,60	0,7
Total	286	229,43	100

Table 10. The complement patterns of *forget* and their frequencies in the 2nd part of the CLMET.

5.2.1 *Forget* + noun phrase

Forget + noun phrase was by far the most common pattern found in the data from the second part of the CLMET with a total of 209 tokens. The percentage of NP complement is roughly the same as in the first part of the CLMET, even though the normed frequency count is notably larger. It must be noted, that the normed frequency count of *forget* in general is also larger in the second part compared to the first part. However, it seems safe to say that NP has a rather stable and established position as a complement that the verb *forget* can take.

As regards the semantic qualities of the NP complement clauses, 205 of the subjects of the clauses were +Human and only four were –Human. In one of the tokens the subject was an animal:

(26) and when they fancied that it [a pigeon] had **forgotten** its former habits, and its old master, they thought that they might venture to employ him nearer home. (Edgeworth 1796-1801, *The Parent's Assistant*)

In the other three tokens with –Human complement the subject was something abstract:

(27) The variegated street shudders at it, for the moment; next moment **forgets** it. (Carlyle 1837, *The French Revolution*)

The NP complement was in 145 of the instances +Abstract and in 64 –Abstract. Out of these 50 employed a personal pronoun or other reference to people:

(28) . . . this faithful servant of the king **forget** his old principles . . . (Lamb 1807, *Tales from Shakespeare*)

Only in 14 cases the NP complement referred to some concrete object:

(29) "' Oh, I **forgot** you' said the tyrant. (Brönte 1847, *Wuthering Heights)*

(30) I won't **forget** my thimble or my scissors. (Edgeworth 1796-1801, *The Parent's Assistant*)

On basis of the study of the present data from CLMET part 2 and the corresponding data from the CLMET part 1, it seems that the subject of a sentence selecting NP complement is most often + Human whereas the NP complement tends to be +Abstract. Furthermore, when it is –Abstract, it is most often a personal pronoun.

To continue with semantics, the senses of the clauses involving NP complements were also analyzed:

OED sense	Frequency
1. 'not to remember or fail to recollect'	88
2. 'to omit or neglect through inadvertence'	27
3. 'to cease or omit to think of'	55
4. 'to disregard or overlook'	36
5. 'to forget oneself'	3

Table 11. The senses of NP complements in CLMET part 2.

As with the first part of the CLMET, the NP complement clauses found among the data in CLMET part 2 appear with a variety of senses. Sense 1 ‘not to remember or fail to recollect’ proved again to be the most common sense associated with NP complement clauses with a total of 88 instances:

(31) Little Arthur had not **forgotten** his father, but thirteen months of absence, during which he had seldom been permitted to hear a word about . . . (Brönte 1848, *The Tenant of Wildefell Hall*)

Second most frequently found sense was sense 3 ‘to cease or omit to think of’ with 55 occurrences while sense 4 ‘to disregard or overlook’ had 36 occurrences:

(32) . . . the circumstance of having got a start by the sudden entrance of her negro servant, and not being able to **forget** him for several hours.” (Hogg 1824, *The Private Memoirs and Confessions of a Justified Sinner*)

(33) . . . although he was little, little child; we shall not quarrel easily among ourselves, and **forget** poor Tiny Tim in doing it. (Dickens 1843, *A Christmas Carol in Prose*)

Sense 2 ‘to omit or neglect through inadvertence’ had 27 occurrences:

(34) I won’t **forget** my thimble or my scissors,” added she, laughing - - “though I used to forget them when I was a giddy girl. (Edgeworth 1796-1801, *The Parent’s Assistant*)

In addition, opposite to the CLMET part 1 findings, also sense 5 ‘to forget oneself’ represented itself in three instances:

(35) Her air was not conciliating, nor was her manner of receiving them such as to make her visitors **forget** their inferior rank. (Austen 1813, *Pride and Prejudice*)

None of them involved a reflexive pronoun, but the nature of sense 5 as given in the OED, i.e. the idea of inappropriate behavior associated strongly with sense 5 in particular, was present. I decided to label them under sense 5 instead of sense 3, which could also have been possible in that the forgetting was momentary. However, the tokens stood out from the rest of the tokens with sense 3 so clearly with

their aspect of inappropriate behavior, that sense 3 seemed less appropriate than sense 5.

As for the extra-semantic aspects involving the NP complements, there were eleven tokens with extraction. Seven of these had evidence of a relativization process and four involved clefting:

(36) an act of kindness$_i$ which$_i$ I could not repay, but never should **forget**. (Brönte 1847, *Agnes Grey*)

(37) He obeyed the call, and in a manner$_i$ that$_i$ I will never **forget**; (Galt 1821, *Annals of the Parish*)

Furthermore, insertion was involved in ten of the tokens with a NP complement:

(38) … to me the stories she had treasured up for her family when once more within the barriers of dear Paris, not **forgetting**, *with that arch,* agreeable vanity peculiar to the French, which they exhibit whilst half ridiculing it . . . (Wollstonecraft 1796, *Letters on Sweden, Norway and Denmark*)

5.2.2 *Forget* + *wh* -clause / *that*-clause

Out of the 286 tokens in the sample 40 had a *that*-clause complement. There were 5 cases where *that* had been omitted, which was relatively less than in the first part of the corpus where *that* was omitted in every fourth token. Then again, the number of *that*-clauses in the first part of the corpus was so small (8 in total) that no extensive generalizations can be made based solely on that. Interestingly, the amount of *that*-clauses exceeds the amount of *to*-infinitives in the second part of the CLMET whereas in the first part of the corpus there were more infinitival complements than *that*-clauses. It will certainly be worth examining further in the coming data to see, whether this is merely a coincidence or is it in fact a proof of ongoing change in the sentential complementation of *forget*.

The subject of the higher and lower clause was the same in only two cases out of the 5 where *that* had been omitted (cf. example 17), whereas in the first

part of the corpus that was the case in all tokens with omitted *that*. Example (39) is an instance where the subjects of the upper and lower clause are different from each other:

(39) But, Mr. Lockwood, I **forget** these tales cannot divert you. (Brönte 1847, *Wuthering Heights*)

Furthermore, of the 35 tokens occurring with *that*-clause complement with *that* not omitted, 10 had the same subject in both higher and lower clause, usually *I* or *you*. Consequently there were 25 tokens left with different subjects in the upper and lower clause and no pattern was to be found as regards the type of subject. Below is an example of such a token:

(40) Do let me **forget** that they are nieces of yours, and let me say that I never saw a more interesting, sweeter pair of sisters . . . (Burns 1780-1796, *The Letters of Robert Burns*)

Regarding the senses of the verb *forget* in connection to the *that*-clause complement clauses, the vast majority of tokens had sense 1:

OED sense	Frequency
1. 'not to remember or fail to recollect'	26
2. 'to omit or neglect through inadvertence'	8
3. 'to cease or omit to think of'	4
4. 'to disregard or overlook'	2
5. 'to forget oneself'	

Table 12. The senses of *that*-clause complements in CLMET part 2.

Out of the tokens with *that* present, 21 had sense 1 as well as all the tokens with *that* omitted, which equals in altogether 26 tokens having sense 1:

(41) Pray, Betty, don't **forget** that Mrs. Strathspey can't breakfast without honey. (Edgeworth 1796-1801, *The Parent's Assistant*)

(42) "You **forget** you have a master here.' says the tyrant. (Brönte 1847, *Wuthering Heights*)

The fact that all the tokens including omission of *that* have the same sense than the majority of tokens with *that* present, is in controversy with Bolinger's generalization

introduced earlier, since according to it a difference in form should equal to a difference in meaning as well, which does not seem to apply to *forget* here.

In addition, sense 2 occurred eight times:

(43) Have you **forgot** that the bracelet—" (Edgeworth 1796-1801, *The Parent's Assistant*)

Sense 3 also had four occurrences:

(44) Joyful congratulations and affectionate greetings between these long-separated parents and their children made them for a while **forget** that Aegeon was yet under sentence of death. (Lamb 1807, *Tales from Shakespeare*)

Furthermore, sense 4 appeared in two tokens:

(45) 'He thinks with his feasts and his wine-cellars to make us **forget** that he is the son of a freedman- and so we will, when we do him the honour of winning his money; (Bulwer-Lytton 1834, *The Last Days of Pompeii*)

Six cases of insertion were found with *that*-clause complements. Only one of them involved omission of *that* which suggests the complexity principle introduced earlier is working as the presence of a complexity factor i.e. insertion leads to using the more complex grammatical variant i.e. retaining rather than omitting *that*. Moreover, there were no insertions to be found with the *wh*-complement sentences.

As for the *wh*-clauses in more detail, there were 8 instances in the second part of the corpus. It seems that the frequency of the *wh*-clauses as complements of *forget* is slowly rising. However, the number of occurrences is rather low considering that it is related to various senses given in the OED and all other secondary sources have a notion of it except Poutsma. Keeping that in mind, it is interesting that in actual data the proportion is perhaps lower than expected on basis of the secondary sources.

All eight tokens conformed to sense 1 'not to remember or fail to recollect':

(46) I shall never **forget** how he spoke for you, when he came running to the embassy to inform us of your arrest. (Borrow 1842, *The Bible in Spain*)

(47) I **forget** whether she used to do this, under your administration. (Austen 1796-1817, *Letters to her Sister Cassandra and Others)*

5.2.3 *Forget* + *to*-infinitive

There were 20 *to*-infinitive complements in the present sample. The number is notably smaller than in the first part of the CLMET; the normed frequency number is 28,62 in CLMET part 1 whereas in the second part of the corpus it is only 16,04. The small amount of infinitives is striking, since it is the only complement offered by all the works cited in section 2 above.

Again, 7 out of the 20 tokens involved verbs of communication. In contrast to the first part of the CLMET, no verb was dominant over the others.

(48) The servants, I suppose, **forgot** to tell you that Mr. Palmer was not in the house. (Austen 1811, *Sense and Sensibility*)

(49) I **forgot** to mention, however, a narrow escape we had, just before we had left an extensive forest on the side of the ... (Marryat 1841, *Masterman Ready*)

Furthermore, all of the sentences with infinitival complement conformed again to the OED sense 2 ‘ to omit or neglect through inadvertence’.

Insertion was found in three tokens:

(50) . . . made some observations on the impropriety of using the word Greek to a British sailor; not **forgetting**, *at the same time*, to speak of the absolute necessity of obedience and discipline on board every ship. (Borrow 1842, *The Bible in Spain*)

However, the presence of insertion does not affect the sense, as all the tokens had the same sense. As regards the infinitival complements, Bolinger’s generalization does not seem to apply very well. In addition, all of the tokens with infinitival

complements conformed to the horror aequi principle as there were no adjacent *to*-infinitives in the data.

5.2.4 *Forget* + *about* + NP

As can be detected from Table 3, Quirk et al, Biber et al. and both dictionaries state that *forget* can be complemented by the *about* + gerund/NP complement. Three tokens were found that had this structure, and they were all of the *forget* + *about* + NP type. All of the instances occurred with the past participle *forgotten*, and each of these three had an intensifier *all* inserted between the matrix verb and the complement. What is more, all of the tokens are in sense 3 ‘to cease or omit to think of’:

(51) and it's very queer to think it, but I've no doubt he has completely **forgotten** all about Ellen Dean, and that he was ever more than all the world to her, and she to him. (Brönte 1847, *Wuthering Heights*)

(52) "And the bananas and the guavas," said Tommy. "Why, we have quite **forgotten** all about them," observed Mrs. Seagrave. (Marryat 1841, *Masterman Ready*)

Based on the first two parts of the CLMET corpus the *about* + gerund/NP complement seems to be on the increase as there were none in the first part of the CLMET whereas the pattern makes an appearance in the second part of the corpus.

5.2.5 Forget + reflexive

Only the dictionaries offer reflexive pronouns as possible complements for *forget* while no grammars studied had a mention of it. There was no evidence of reflexive pronouns used as complements of *forget* in the first part of the CLMET, either.

However, 4 tokens with reflexive pronoun complements were found in the data from the second part of the CLMET.

According to both dictionaries the pattern *forget* + reflexive conveys an idea of improper behavior of some sort, and it seems that all four instances have such implications. The OED sense 5 in its original form is 'to lose remembrance of one's own station, position, or character; to lose sight of the requirements of dignity, propriety, or decorum; to behave unbecomingly' and all the tokens conform to that:

> (53) her soul was calm--the tempest had subsided--and nothing remained but an eager longing to **forget** herself--to fly from the anguish she endured to escape from thought--from this hell of disappointment. (Wollstonecraft 1798, *Maria*)

The reflexive is often used idiomatically, as in the above example. However, there are also cases where the reflexive is used in a more literal way, as can be detected from the example below:

> (54) ... took Tiny Tim beside him in a tiny corner at the table; the two young Cratchits set chairs for everybody, not **forgetting** themselves, and mounting guard upon their posts, crammed spoons into their mouths, lest they should shriek ... (Dickens 1843, *A Christmas Carol in Prose*)

Example (54) could be analyzed as having some other sense as well, but also so that the young Cratchits did not forget themselves in that they did not lose remembrance of their station, which justifies their having chairs for themselves as well as for all the other people by the table.

5.2.6 *Forget* + gerund

There were only two tokens with a gerund complement in the corpus data. These two are very similar in many ways. Both come from the same text and there is a modal auxiliary *shall* in the matrix clause:

> (55) I never shall **forget** conning over the Catalogue which a friend lent me just before I set out. (Hazlitt 1821-22, *Table Talk*)

(56) I shall not easily **forget** bringing him my account of her first appearance in the Beggar's Opera. (Hazlitt 1821-22, *Table Talk*)

The gerund was not noted by the OED whereas Collins/Cobuild associated it with sense 1, which is the appropriate sense in these example sentences as well. As was stated in section 3.2, *-ing*-forms typically refer to something more concrete whereas the *to*-infinitive is used in reference to something hypothetical (Bolinger 1968, 123-124). That holds true for these sentences as well, since the gerundial in both instances refers to something that has actually happened. As for the extra-semantic aspects, no structural discontinuity or extra-semantic constraints were present in the data.

It is curious that there are so few gerunds in the corpus data, since all the grammars and dictionaries state that *forget* can take *-ing*-clause complements. However, there were no examples of *forget* + *about* + *-ing*-clause although Quirk et al (1985, 1193) claim that the *-ing*-clause is rare without preceding *about*. Since there was no evidence of *-ing*-clauses in the data collected from the first part of CLMET it could be that the gerund was slowly gaining ground. Furthermore, according to the theory of the Great Complement Shift, gerunds should become established at the expense of infinitival forms. The present results seem to be in contrast with that theory.

5.3 CLMET part 3

The third part of the CLMET contains 3,982,264 words in 52 texts from 1850-1920. The query for *forget* and its inflectional forms yielded a total of 994 hits and after sorting a sample there were 334 tokens left for further examination. Out of the 334 tokens 66 proved to be irrelevant; there were 27 instances involving a zero complement and in 39 cases *forget* was used as an adjective:

(57) “By heaven, I **forgot**!” cried the king. (Hope 1898, *Rupert of Hentzau*)

(58) . . . and wondered whose was the beauty that it had upborne through the pomp and pageantry of a **forgotten** civilization --- (Haggard 1887, *She*)

Consequently, 271 tokens were then taken into account when analyzing the data.

Frequencies can be detected in the table below:

Type of complement	**Number of tokens**	**Normed frequency**	**Percentage**
NP	172	129,57	64,2
That-clause	48	36,10	17,9
To-infinitive	22	16,57	8,2
Wh-clause	10	7,53	3,7
About + NP	10	7,53	3,7
Reflexive	6	1,51	2,2
Total	268	201,90	100

Table 13. The complement patterns of *forget* and their frequencies in the 3rd part of the CLMET.

5.3.1 *Forget* + noun phrase

As was the case with the first two parts of the CLMET, the NP was by far the most common complement in the third part of the corpus as well. However, as can be detected by the normalized frequencies, its frequency is somewhat smaller than in the second part of the CLMET even though it is still larger than in the first part of the corpus. Interestingly the percentage of NP complements out of all NP complements is smaller in CLMET part 3 than in the two previous parts. In all, there were 172 NP complement clauses in the data from the CLMET part 3.

As regards the semantic qualities of the NP complement clauses, the subject of the clause was again categorized either +Human or –Human and the NP complement +Abstract or –Abstract. The subject of the upper clause was +Human in 170 of the cases whereas there were only 2 –Human subjects: ‘a day’ and ‘God’. The NP complement was +Abstract in 114 tokens and thus 58 tokens involved a –Abstract

NP complement. 28 of the –Abstract NPs were +Human; in most cases a personal pronoun.

(59) But anybody can **forget** an umbrella, as anybody might forget a shed that he has stood up in out of the rain. (Chesterton 1910, *What's Wrong With the World*)

(60) 'Divorce her—turn her out! She has **forgotten** you. Forget her!' (Galsworthy 1906, *The Man of Property*)

The distribution of the OED senses was rather similar compared to the earlier parts of the CLMET in that most senses are widely represented, as can be detected from table 14 below:

OED sense	Frequency
1. 'not to remember or fail to recollect'	69
2. 'to omit or neglect through inadvertence'	46
3. 'to cease or omit to think of'	36
4. 'to disregard or overlook'	14
5. 'to forget oneself'	7

Table 14. The senses of NP complements in CLMET part 3.

The most common was again sense 1 with 69 instances:

(61) "How are my proteges?" asked Cecil, who took no real interest in them, and had long since **forgotten** his resolution to bring them to Windy Corner for educational purposes. (Forster 1908, *A Room with a View*)

Sense 2 appeared 46 times in the data, including the cases with leaving something behind as well as omitting to mention or leaving something unnoted:

(62) Myself I had **forgotten** my cigar-case. (Jerome 1909, *They and I*)

(63) "Did she not mention a companion and friend -- a woman named Mrs. Clements?" "Oh yes! yes! I **forgot** that. She told me Mrs. Clements wanted sadly to go with her to the lake and take care of her, and . . . (Collins 1859-60, *The Woman in White*)

Furthermore, sense 3 had 36 occurrences whereas sense 4 was discernible in altogether 14 instances:

(64) Don't altogether **forget** poor Fanny in thinking of me. (Collins 1859-60, *The Woman in White*)

(65) Dear Annie appeared to believe that it could; saying that if the rich continually chose to **forget** the poor, a man who forced them to remember, and so to do good to themselves and to others, was a . . . (Blackmore 1869, *Lorna Doone, A Romance of Exmoor*)

In addition, sense 5 also appeared seven times:

(66) **Forgetting** his usual self-control and the forms of public utterance, he broke out into a long and abusive harangue. (Churchill 1899, *The River War, an Account of the Reconquest of the Sudan*)

(67) I shook of Sergeant Cuff's arm, and, **forgetting** my manners, pushed him through the door to make my own inquiries for myself. (Collins 1868,*The Moonstone*)

When considering the extra-semantic elements detected amongst the data, structural discontinuity was analyzed in 21 tokens. There were traces of extraction processes discernible in nine tokens. The most common extraction structure was relativization with five instances in the data:

(68) Her ladyship, the Countess, explained her unexpected appearance at the inn by telling Fanny that she had come to bring one or two little messages$_i$ which$_i$ Miss Halcombe in her hurry **had forgotten**. (Collins 1859-60, *The Woman in White*)

In addition, there were two cases of clefting as well as two cases of topicalization:

(69) It is altogether a place$_i$ that you won't forget, a place to open a man's soul, and make him a prophesy . . . (Hughes 1857, *Tom Brown's Schooldays*)

(70) The sense of desperation$_i$ which this incident produced I shall not easily **forget**. (Gosse 1907, *Father and Son, A Study of Two Temperaments*)

Furthermore, in 4 sentences there was an element inserted:

(71) . . . with his Lord Feltre, bound to make an inspection of Syrian monasteries, and **forget**, *if he could*, the face of all faces, another's possession by the law. (Meredith 1895, *The Amazing Marriage*)

5.3.2 *Forget* + *wh*-clause / *that*-clause

There were altogether 48 *that*-clauses among the data and ten instances of *wh*-clause complement. The frequency of the *wh*-clauses as well as of the *that*-clauses was on the rise in the third part of the corpus compared to the earlier two parts. This suggests that the change in progress suspected in subsection 5.2.2 is in fact happening as the amount of *that*-clauses is steadily increasing throughout the three consecutive parts of the CLMET. The amount of the *wh*-clauses is also on the rise based on the normed frequency numbers, but even though there is clearly a change in progress, it is much more subtle than with the *that*-clauses.

The subject of the upper clause in *that*-clauses was in seven out of nine cases the first person personal pronoun *I*. The other two tokens had *you* and *he*. The subject of the lower clause was also a personal pronoun in seven of the cases whereas *it* was present in two tokens. The same personal pronoun was used only in one token in both the upper and the lower clause:

> (72) "I was **forgetting** I asked him to come and have lunch with us, early. . . . (Bennett 1908, *The Old Wives' Tale*)

That was omitted in nine of the tokens. Insertion was found in 9 of the tokens with *that*-clause complement in two of which *that* was omitted. Rohdenburgs's complexity principle seems to be effective as most of the tokens involving insertion have retained *that*. Furthermore, there was one instance of a *wh*-clause with insertion. The inserted element was an adjunct (e.g. *in the flurry of the moment*), a subordinate clause or e.g. *sir* or *dear Christina*, but no variant was dominant over the others:

> (73) I was thinking of other things, and **forgot**, *as any one might easily understand*, that I was steering, and the consequence was that we had got mixed up . . . (Jerome 1889, *Three Men in a Boat*)

> (74) The effect of time was such that even Mr. Critchlow appeared to have **forgotten** *even* that she had been indirectly responsible for her farther's death. (Bennett 1908, *The Old Wives' Tale*)

(75) . . . he exclaimed angrily; “you **forget**, *my dear Christina*, that you have to deal with a stomach that is totally disorganized,” . . . (Butler 1903, *The Way of All Flesh*)

The two in which *that* was omitted were almost similar:

(76) I **forgot**, *of course*, you don’t know. (Bennett 1902, *The Grand Babylon Hotel*)

(77) Oh, I **forgot**; *of course*, you knew her before. (Forster 1908, *A Room with a View*)

As to the semantic side of the *wh*-complement clauses, nine out of the ten instances conformed to sense 1 ‘not to remember or fail to recollect’:

(78) It is so easy to talk of “passing emotion,” and to **forget** how vivid the emotion was ere it passed. (Forster 1910, *Howards End*)

In addition to that, one token represented sense 3 ‘to cease or omit to think of’:

(79) No, I cannot **forget** how they behaved at Mr. Eager’s lecture at Santa Croce. Oh, poor Miss Honeychurch! (Forster 1908, *A Room with a View*)

The distribution of different senses that *forget* is said to have is much wider among the tokens with *that*-clause complements. Table 15 introduces the figures based on the actual findings in the present data:

OED sense	Frequency
1. 'not to remember or fail to recollect’	20
2. 'to omit or neglect through inadvertence'	11
3. 'to cease or omit to think of'	13
4. 'to disregard or overlook'	4
5. 'to forget oneself'	

Table 15. The senses of *that*-clause complements in CLMET part 3.

Again, in violation to Bolinger’s generalization, the tokens involving omission of *that* were no different than the ones that had retained *that* as regards the sense. In the data from the CLMET part 3 there are 20 tokens with sense 1 ‘not to remember or fail to recollect’:

(80) . . . we are apt to **forget** that offspring is only a full-sized reproduction of the parent-- … (Butler 1880, *Unconscious Memory*)

In addition there were 13 tokens with sense 3 ‘to cease or omit to think of’ and 11 tokens with sense 2 ‘to omit or neglect through inadvertence’:

(81) Go now, Holly, go, and, if thou canst, try to **forget** that thou hast of thy folly looked upon Ayesha’s beauty,” (Haggard 1887, *She*)

(82) Mrs. Furze made a dash at her husband’s clay pipe, **forgetting** that its destruction would not make matters better; but she only succeeded in upsetting the chair . . . (Rutherford 1893, *Catherine Furze*)

Furthermore, four tokens occurred with sense 4 ‘to disregard or overlook’:

(83) I cannot have you speaking out of your role. It makes my poor head go round, and I think you **forget** that I am very ill.” “Your head won’t go round if only you’ll listen to my argument,” said Margaret. (Forster 1910, *Howards End*)

The tokens with *that* omitted had all either sense 1 or sense 2:

(84) But I’m **forgetting** you are King, sire. (Hope 1894, *The Prisoner of Zenda*)

(85) I was **forgetting** I asked him to come and have lunch with us, early. (Bennett 1908, *The Old Wives’ Tale*)

All in all, the frequency and use of *that*-clause complements seems to be rather stable on basis of all three parts of the CLMET.

5.3.3 *Forget* + *to*-infinitive

With 22 occurrences in the present data drawn from the CLMET part 3, the number of *to*-infinitive complements was somewhat consistent with the results from the second part of the CLMET. It seems that the frequency of infinitival complements has decreased notably from what it was in the early 18th century, but has remained about the same since the late 18th century until the early 20th century. The BNC data will be particularly useful to observe how the situation has developed later on and whether there is still a change in progress. As with the earlier data from the CLMET parts 1 and 2, all of the present tokens had the OED sense 2 ‘to omit or neglect through inadvertence’ or sub-sense 2d ‘to omit mention or leave unnoted’. Furthermore, the

proportion of communicative verbs in the lower clause was lower in the third part of the CLMET: only six instances were found. In three of the cases the verb was *to ask*, twice *to tell* and once *to insist*:

> (86) I forgot to tell you of an appointment of mine at a place called Canleys, about twenty miles or more from here. (Meredith 1895, *The Amazing Marriage*)

Among the other than communicative verbs no clear tendency to favor a certain type of verbs was to be found.

There were no instances of extraction but instead two instances with an inserted element were found:

> (87) . . . fety in my breast to which I had for some days been a stranger, I crept off to my own little sepulchre, not **forgetting** *before I laid down in it* to thank Providence from the bottom of my heart that it was not a sepulchre indeed, . . . (Haggard 1887, *She*)

> (88) . . . the first Arab historian who shall investigate the early annals of that new nation will not **forget**, *foremost among the heroes of his race*, to write the name of Mohammed Ahmed. (Churchill 1899, *The River War, An Account of the Reconquest of the Sudan*)

However, the sense remains the same, so Bolinger's generalization does not apply here, either. In addition to insertion, there were two tokens that violated against the horror aequi principle:

> (89) Spent the evening packing, Carrie told me not *to* **forget** to borrow Mr. Higgsworth's telescope, which he always lends me, knowing I know how to take care of it. (Grossmith 1894, *The Diary of Nobody*)

5.3.4 *Forget* + *about* + NP

The pattern *about* + NP followed *forget* in ten of the tokens. The possibility of an *about* + gerund complement suggested by Quirk et al. and Biber et al. (cf. e.g. Table 3) did not reveal itself in the third part of the CLMET any more than it did in the second part of the CLMET. In four out of the ten cases there was insertion, which in all of the cases was the intensifying *all*:

> (90) I shall doze there for a fortnight, and **forget** *all* about the "so-

called" nineteenth century." (Gissing 1893, *The Odd Woman*)

Furthermore, all of the tokens with the *about* + NP pattern had again the OED sense 3 'to cease or omit to think of, let slip out of the mind, leave out of sight, take no note of' although the complement pattern under discussion here is suggested to appear rather with OED sub-sense 1d 'not to recall the facts concerning; not to remember to take action in the matter of'. In the ten tokens found in the third part of the CLMET it is arguably not about 'not recalling the facts concerning' but rather about 'ceasing or omitting to think of, let slip out of the mind or take no note of'. For example in the following example sentence (91) the *forgetting* has the sense 'omit to think' or 'let slip out of the mind' rather than not remembering 'the facts concerning':

(91) 'I don't see that the review is such a terrible one, after all. Besides, everybody has **forgotten** about it by this time. I'm sure the opening is good enough for any book ever written. . . . (Hardy 1873, *A Pair of Blue Eyes*)

5.3.5 *Forget* + reflexive

The reflexive complement was present in 6 tokens. Comparing the percentages it seems to be somewhat on the increase, especially when considering that there were no instances of the reflexive in the first part of the corpus. The reflexives in the second part of the CLMET as well as in the third part of the corpus seem to conform well to the OED sense 5. The reflexive pronoun always matched with the subject:

(92) "My dear, you are **forgetting** yourself." (Forster 1910, *Howard's End*)

(93) On hearing those words, the infernal detective-fever began, I suppose, to burn in me again. At any rate, I **forgot** myself in the interest of guessing this new riddle. (Collins 1868, *The Moonstone*)

Opposite to CLMET part 2 findings in section 5.2.5, all tokens were used idiomatically.

5.4 Summary of CLMET results

After looking into the data extracted from the first two parts of the CLMET corpus a brief summary is appropriate before moving on to the BNC data. Table 16 below shows the distribution of complements in all parts of the corpus:

Type of complement	**CLMET part 1 Raw frequency**	**Normalized frequency**	**CLMET part 2 Raw frequency**	**Normalized frequency**	**CLMET part 3 Raw frequency**	**Normalized frequency**
Noun phrase	75	107,33	209	167,66	172	129,57
That –clause	8	11,45	40	32,09	48	36,1
To –infinitive	20	28,62	20	16,04	22	16,57
Wh –clause	4	5,72	8	6,42	10	7,53
About + NP			3	2,41	10	7,53
Reflexive			4	3,21	6	4,52
Gerund			2	1,6		
Total	**107**	**153,12**	**286**	**229,43**	**268**	**201,9**

Table 16. The distribution of complements in CLMET part 1, 2 and 3.

As can be seen from the table, there are both similarities and differences discernible. Firstly, the first part had examples of only four types of complements, whereas in the later parts the variety was much greater, which could simply be due to the larger size or alternatively it could refer to a change taking place over the 18th, 19th and 20th centuries. Indisputably the most common complement in all parts of the corpus was the non-sentential NP complement. However, the normalized frequencies suggest that the frequency of the NPs is decreasing and at the same time the frequencies of all other complements except the gerund are on the increase. That would imply that in general the use of complements of the verb *forget* is becoming more diverse over time and the proportion of sentential complements is increasing. It will be interesting to see whether the results from the BNC corroborate with the hypothesis.

The reflexive was conspicuously absent in the first part of the CLMET and only 4 were found in the second and 6 in the third part, even though intuitively it comes across as somewhat archaic and I would have expected it to emerge more

forcibly among the texts from the early and mid-18th century. Another prospering complement pattern is *forget* + *about* + NP, which is clearly gaining ground.

Sentential complements were still in the minority, although all of the sentential complement variants were gaining ground in the last part of the CLMET. *That*-clause complement seems to be clearly advancing along with the *wh*-clause, although the change there is not as drastic. There were no instances of the gerund in the first or the third part of the CLMET, in addition to which only two were found in the second part. The amount of *to*-infinitives is in part 2 lower than in part 1, but interestingly the frequency does not decrease in part 3 as could be expected but rather it is slightly on the increase. According to the Great Complement Shift theory the gerund should be increasing and the *to*-infinitive decreasing, but with *forget* the theory does not seem to apply very effectively and the possible effects are more obscure.

6 *Forget* in the British National Corpus

The British National Corpus (the BNC) consists of about a hundred million words, and it consists of multiple text domains. The query for *forget* and its inflectional forms was restricted to text domains of *Imaginative prose* and *Informative: Natural and pure sciences* and *Informative: World affairs* so as to offer a sample more comparable to that of the CLMET texts, since as pointed out earlier in 2.2, the data in the CLMET is from literary and formal texts (de Smet 2005, 71). The selected domains contain altogether roughly about 37,500,000 words from 1,105 texts and the query for *forget* in those domains of the BNC yielded altogether 5742 hits in 773 different texts. The number of hits in the BNC was so high that I chose to use the random selection tool available online in the BNC web page to keep the sample manageable. Thus 957 randomly selected tokens were left for further analysis. Out of these 124 proved irrelevant for the present thesis leaving 833 tokens for further analysis. Most of the irrelevant tokens, 97 tokens, had a zero complement, which I have chosen to omit from the thesis.

(1) **BPO 2774** Katelina said, "I hadn't **forgotten**.'

Furthermore, there were 25 cases where *forget* was used as an adjective and one instance of a nominal use of *forgetting*:

(2) **HGL 840** Winchell stand eager and **forgotten** several feet away.

(3) **CAF 1206** . . . the struggle of humanity against power is the struggle of memory against **forgetting**, as Kundera puts it . . .

In addition, there was a misspelling of 'forget-me-nots', which was of course deemed irrelevant for this study as well:

(4) **J2F 384** The thing here was that all the flowers were blue — delphiniums and cornflowers and **forget** me nots round a sundial in the middle of a smooth green velvet lawn.

Even though precision is notably weaker than with the CLMET data, there are enough tokens left for analysis. Table 17 below introduces the frequencies of complements found in the BNC:

Type of complement	**Number of tokens**	**Normed frequency**	**Percentage**
Noun phrase	497	79,4	60
That-clause	126	20,13	15,1
About + NP	84	13,42	10
To-infinitive	62	9,9	7,4
Wh-clause	51	8,15	6,1
Reflexive	8	1,28	1
Gerund	3	0,48	0,4
About + V-*ing*	2	0,32	0,2
Total	**833**	**133,01**	**100**

Table 17. The complement patterns of *forget* and their frequencies in BNC.

6.1 *Forget* + noun phrase

The most common complement of *forget* in the BNC sample was the NP, which is well in accordance with the CLMET results. By considering the normed frequency counts from all parts of the CLMET as well as from the BNC it could be claimed that the most common complement of *forget* is indisputably the NP complement and its position seems rather established. Altogether 60% of post-head complements of *forget* found in the BNC were NPs. However, it must be noted that the normed frequency count is considerably lower in the BNC data than with the CLMET data.

To continue with the present data from the BNC, there were 497 NP complements in the sample. The semantic qualities were again analyzed and the subject of the clause was marked +/- HUMAN and the object +/- ABSTRACT. It now seems safe to say that the subject of a *forget* + NP construction is most often +HUMAN, since there were only five instances of -HUMAN subjects in the present sample, i.e. *obituaries, countries, steps, our system* and *the reverse of a system*:

(4) **FNR 577** Almost all monotonic operators are reversible: If the operator adds a new truth, then its reverse just **forgets** that truth again.

As regards the nature of the NP complements then, 318 were +ABSTRACT and 179 -ABSTRACT, out of which 104 proved to be +HUMAN:

(5) **GW0 2918** 'I'll never **forget** you,' Lucy said.

There was no clear pattern or tendency to be noted among -ABSTRACT -HUMAN NP complements as they seemed rather random. Based on the analyses of the semantic qualities of NP complement clauses throughout the Late Modern English period one could claim that *forget* selects foremost +HUMAN subjects while the nature of the NP complement tends to be +ABSTRACT, and in the case of -ABSTRACT objects, it tends to be a personal pronoun rather than -HUMAN in quality.

To continue with semantics, the different senses of *forget* were again all represented in the data:

OED sense	Frequency
1. 'not to remember or fail to recollect'	185
2. 'to omit or neglect through inadvertence'	109
3. 'to cease or omit to think of'	132
4. 'to disregard or overlook'	67
5. 'to forget oneself'	4

Table 18. The senses of the NP complements in the BNC.

The most common one was sense 1 with 185 hits, which also corroborates with the CLMET findings:

(6) **HH5 401** I never **forget** good meals. I mean, if you have starved like I have in the wilds of Muscovy or the deserts of North Africa, you always remember what you have eaten.

There were only two cases of insertion among the tokens conforming to sense 1. Furthermore, extraction was found in 22 tokens. 17 of these were cases of relativization:

(7) **ANR 388** The crowd cheered the Empress who responded with a curtsy — that famous gesture$_i$ which$_i$ those who witnessed it never **forgot** t_i and which became the hallmark of Eug&#65533;nie's graciousness.

There were also four instances of clefting and one sentence with evidence of a topicalization process:

(8) **JYB 990** It isn't something$_i$ you easily **forget** t_i'.' Silas spoke quietly.

(9) **G13 2096** A face$_i$ I have never **forgotten** t_i.

Interestingly sense 1 tends to attract extractions and especially relativization, or vice versa: when relativization has taken place, the sense is typically 1.

Second most common sense was sense 3 with 132 occurrences:

(10) **AD0 128** You may as well **forget** the idea of banning things.

It is noteworthy, that most of the cases with insertion were in sense 3. In fact, out of the ten cases of insertion five were in this particular sense, whereas the other half was divided between senses 1, 2 and 4. Furthermore, according to the corpus data when the element inserted is the intensifier *all*, the appropriate sense is sense 3:

(11) **HA9 2672** 'Are you about to **forget** *all* your fine words and simply take what you want?

All the insertions in sense 3 were in fact *all* while no such insertions were found in any of the other senses. Moreover, altogether 109 tokens had sense 2:

(12) **HH1 3970** Don't **forget** a cloak, Eleanor.

Two of the tokens involved insertion and in two cases extraction (i.e. relativization) was found:

(13) **H0R 2185** worried at the gaps in her education, anxious about nuns and antique dealers, she had **forgotten** *for some time* the necessity for personal happiness.

(14) **ABW 1237** A girl run out with his buns$_i$ which he'd **forgotten** t_i.

Furthermore, there were 67 tokens in sense 4:

(15) **AJD 139** ‘For the fourth time, they [the Labour party] are coming back to the people and asking everybody to forget two recessions, **forget** millions unemployed, forget thousands of bankruptcies and thousands more whose homes have been repossessed.

As to complexity factors, there was insertion in one token and relativization in three tokens:

(16) **G1D 1580** ‘Can we afford it?’ asked Scarlet, **forgetting** *on the instant* her vow to abandon the practice of hope.

(17) **AD2 707** To connect these two points, I would stress the obvious point_i which may be **forgotten** t_i when one simply looks at the fortunes of particular party organizations.

What is more, there were also four tokens with NP complements used in sense 5:

(18) G17 1707 Outside a heavy mist obscured everything, and the departing Khans and Noyons went accompanied by lantern bearers and pairs of Merkut guards against the interference of revellers who might **forget** their manners.

The NP complement referred to inappropriate or unbecoming behavior in all of the four tokens which were analyzed as having sense 5. The reference to the unexpected behavior was so evident that the tokens fit the criteria set by the OED for sense 5 better than the criteria for the other senses.

To summarize, NP complement is the most frequent choice of complement and the most typical sense is sense 1. When there is insertion involved in the clause, it can have almost any of the senses, unless the inserted element is *all*; in that case it is most probably sense 3. When there is extraction involved, it tends to be relativization and the most frequent sense is sense 1, although it can also be some other sense as well.

6.2 *Forget* + *that*-clause

There were altogether 126 *that*-clause complements in the BNC sample. In 49 cases *that* was omitted. There was insertion in 10 tokens with *that* visible, and in 6 tokens with *that* omitted i.e. there was insertion in altogether 16 tokens with *that*-clause complement.

The distribution of senses in *that*-clauses differed somewhat from the CLMET data results. Table 19 shows the situation with current data:

OED sense	Frequency
1. 'not to remember or fail to recollect'	55
2. 'to omit or neglect through inadvertence'	62
3. 'to cease or omit to think of'	5
4. 'to disregard or overlook'	5
5. 'to forget oneself'	

Table 19. The senses of *that*-clause complements in the BNC.

The frequencies of senses 1 and 2 were about the same in both the cases with *that* omitted and retained, which suggests that they are at least to an extent mutually interchangeable. Of the tokens with *that* present, 32 were in sense 1 (example 19) and 36 in sense 2 (example 20):

(19) **HGT 4278** There was a terrible urge within her to just run away and hide, curl up into a tiny ball and **forget** that she had made such a stupid, stupid mistake by allowing herself to fall in love with someone as ruthless and cold as Luke Denner.

(20**) H7W 2111** So much of my time is spent designing or in meetings it's all too easy to **forget** that this,' he gazed around him, 'is the purpose of it all.'

With *that* omitted there were 23 tokens in sense 1 (example 21) and 26 in sense 2 (example 22):

(21) **A1Y 238** Supporting the leadership, Garfield Davies, general secretary of the shop workers' union Usdaw, said delegates should not **forget** the objective was to create a nuclear-free world, not just a nuclear-free Britain.

(22) **BMU 1927** ‘As a matter of fact, I'd **forgotten** you were coming — and we couldn't have had tea with you in any case, because Susan's ill.

As for the tokens with *that* and an inserted element, 5 were in sense 1 and 4 in sense 2. With the tokens with *that* omitted insertion seemed to affect the sense more, as there was only one token in sense 1 and five in sense 2. However, the amount of such cases in the present sample is unfortunately so low that a larger sample would be needed to see whether it is a mere coincidence or an actual tendency to favor sense 2 when *that* is omitted and there is insertion involved.

In addition to senses 1 and 2, there were also 5 instances of sense 3 (example 23) and 5 instances of sense 4 (example 24), in all of which *that* was present:

(23) **JYC 2148** Today, she decided, she would **forget** that she couldn't remember.

(24) **C8F 127** If, as indicated above, there are paramount safety and environmental reasons for addressing the traffic problem in residential areas, then it should not be **forgotten** that there are social arguments too.

To summarize, the BNC data suggests that if *that* is omitted, the selection of senses is restricted to senses 1 or 2 whereas the sense is most often 1 or 2 in cases where *that* is retained, but it can also be 3 or 4. Insertion seems to be possible mostly with sentences involving sense 1 or 2.

What is more, there was also evidence of extractions in the data. There were in fact two cases of clefting in *that*-clauses:

(25) **ACE 2124** What we’re **forgetting** is that you we’re an actor in civilian life.

(26) **HWE 2227** What you seem to **forget** is that it’s on sixteen years since you left this house.

6.3 *Forget* + *about* + NP / V-*ing*

There were 84 instances of the complementation pattern *forget* + *about* + NP in the BNC data. Furthermore, there were also two instances of the so far nonexistent pattern *forget* + *about* + V-*ing*. The patterns involving *about* are clearly on the increase based on the comparison of the normed frequency counts, while the normed frequency of the *about* + NP complement in the BNC is 13,42 and in the CLMET the normed frequencies were 2,41 in the second part and 7,53 in the third (cf. table 16). There were no instances of the pattern in the first part of the CLMET. Furthermore, the pattern *about* + NP has been clearly increasing throughout the corpus data used in this thesis, whereas the pattern involving the gerund form has been absent in the previous data from the CLMET and appears only in the BNC data, and even there only twice. All of this indicates that the complement involving *about* is definitively becoming more common.

To recall what the secondary sources said about the complements involving *about* + NP / V-*ing*, only Quirk et al (1985, 1193) together with Biber et al (1999, 742) acknowledged the pattern *forget* + *about* + V-*ing* whereas the pattern *forget* + NP was absent from all the grammar books and only emerged in the dictionaries i.e. the OED and Collins/Cobuild. Based on the information derived from secondary sources it is perhaps not surprising that *about* + NP has not been such a common complement chosen by *forget*, and consequently the obvious rise in frequency counts throughout the Late Modern English period is intriguing.

As to the pattern involving the gerund, Quirk et al also state that the pattern *forget* + gerund is rare without preceding *about* (ibid, 1193). The verifiability of the claim remains somewhat unclear as there are so few gerundial complements in the data in general, however it is worth noting that there were none or only a few gerunds in the CLMET data and no tokens at all with the pattern *forget* + *about* + V-

ing, which could suggest that the claim does not apply to the authentic usage of *forget*. Furthermore, the pattern mentioned above appears only in the BNC and the frequency is still miniscule. As a conclusion it could be claimed that *forget* does not seem to prefer gerundial complements at all, and when there is a gerundial complement in a sentence it may be with or without preceding *about*. The presence of *about* in a gerundial complement apparently depends on the sense in which *forget* is used in a sentence. Both of the tokens in the BNC data involving the pattern *about* + V-*ing* were namely in sense 3 whereas the other gerundial tokens in the CLMET and in the BNC were mostly in sense 1. As noted earlier, the number of gerundial complements is low, but although no extensive rules can be made on basis of such low frequency counts, some clear tendencies should still be noted. Firstly, it seems that when there is a gerund complement with an auxiliary verb, the appropriate sense is sense 1. Secondly, when the preceding *about* is present, the appropriate sense is 3.

Finally, below are the tokens involving the pattern under discussion above:

(27) **FSF 3074** He might have **forgotten** about needing a visa because he'd developed the prisoner mentality, so common among people who were never permitted to travel.

(28) **GUF 1855** The events of the last few days had made Charles **forget** about Jacqui's flat being done over, but inside it the evidence was all too clear.

To turn the focus towards the tokens with the pattern *forget* + *about* + NP, as noted earlier in the beginning of the present section, the frequency of the pattern is higher than with the CLMET data. All of the tokens in the CLMET involving the pattern had sense 3 whereas there is a little more diversity in the BNC data:

OED sense	Frequency
1. 'not to remember or fail to recollect'	16
3. 'to cease or omit to think of'	67

Table 21. The senses of *about* + NP complement in the BNC.

Thus it seems that as the pattern is becoming more common, there is also a change in progress as regards the choice of sense. Having said that, sense 3 was still the prevalent choice with altogether 67 instances:

(29) **HJ3 1266** 'My job is to go on stage and make people **forget** about their problems -- …

In addition though, there were 16 instances of sense 1:

(30) **HHA 3362** 'Last night,' he snapped, 'was entirely my fault. Night-time madness isn't appealing, seen in the cold light of day. I'm not proud of myself, and I want to **forget** *all* about it.

There was nothing in particular that would have revealed the secret behind the choice of sense, rather it seemed rather random.

What is more, there was an inserted element in 26 of the tokens with the pattern *forget* + *about* + NP. Only three tokens had an adjunct inserted (*momentarily*, *more* and *entirely*) while in 23 cases the inserted element was the intensifier *all*:

(31) **AN7 3413** 'Just the drink talking. He'll have **forgotten** *all* about it in the morning,' Maggie smiled in reply.

However, it seems that the presence of an inserted element does not affect the sense as the senses of the tokens with insertion divided rather equally between senses 1 and 3

Furthermore, there were also three cases of extractions whereas in the CLMET data there was no evidence of such structural discontinuity. This, too, adds to the earlier claim about the change in progress as regards the use of complements involving *about*. There was one instance of topicalization, one of relativization and one of clefting:

(32) **GUK 1661** A language she once knew$_i$ but had **forgotten** about t_i.

(33) **HNR 597** . . . pretending to be looking at the waterlilies and what

might be a nest of some kind$_i$, over there, which$_i$ two who are clamped together mouth to mouth have **forgotten** about t$_i$.

(34) **J17 559** An interest in music$_i$ that I had rather **forgotten** about t$_i$ until the thought of Millie in the lane reminded me.

6.4 *Forget* + *to*-infinitive

The infinitival complement was present in 62 of the tokens. According to the Great Complement Shift the *-ing*-forms should be expanding at the expense of *to*-infinitives, and so far the case with *forget* has been complicated. The proportion of infinitival complements began to decrease only to become slightly more common again, and then decrease again. The development has not been straightforward, but the frequency is still on a slow decline. The normed frequencies in the BNC data (cf. tables 16 and 17) suggest that the proportion of *to*-infinitives as well as the proportion of *that*-clauses is presently on decrease. Even though the amount of gerunds is perhaps not increasing notably, at least there is a change to be noted in the frequency of the *–ing*-forms. Therefore it could be claimed, that the Great Complement Shift is slowly beginning to affect the complementation system of *forget* as well, even though the changes are still relatively subtle.

Again, in concordance with earlier results, all of the example sentences conformed to sense 2 ‘to omit or neglect through inadvertence’ or 2b ‘to omit to take or leave behind’ or 2d ‘to omit mention or leave unnoted’. The proportion of verbs of communication in the lower clause seemed to be on the increase, as in 21 cases the lower verb was e.g. *to tell*, *to ask*, *to say*, *to mention* and even *to lisp*:

(35) **BP7 678** It’s typical of him to **forget** to tell me he’s entertaining.

(36) **EFW 1807** “There was something I **forgot** to mention to you earlier,”

There was no evidence of insertion or extractions, but there was one token with two adjacent infinitives, which is not in accordance with the horror aequi principle:

(37) **BP7 678** It's typical of him *to* **forget** to tell me he's entertaining.

There were no other violations against the horror aequi principle in the BNC data and in the CLMET there were no more than two, either, which suggests that in general the principle tends to be effective with *forget*.

6.5 *Forget* + *wh*-clause

As *that*-clauses have advanced so notably, they were treated separately unlike in the CLMET chapter. Therefore *wh*-clauses are also discussed independently in this present chapter. There were 51 *wh*-clause complements in the BNC data. The proportion of *wh*-clause complements seems to be on a slow, but steady increase.

As regards the semantic aspect, the majority of the tokens had sense 1, as can be detected from table 22:

OED sense	Frequency
1. 'not to remember or fail to recollect'	38
3. 'to cease or omit to think of'	5
4. 'to disregard or overlook'	8

Table 22. The senses of *wh*-clause complements in the BNC.

In the CLMET data nearly all instances of *wh*-clause complements conformed to sense 1, which is the predominant choice in the BNC as well:

(38) **FS3 230** I **forget** who I am!

However, it is worth noting that as the frequency of the *wh*-complements is rising, also the senses used in connection to it are becoming less restricted. There were already some instances of sense 3 usage in the CLMET data, but in the BNC the amount of tokens in other senses than sense 1 as well as the variety of senses themselves was larger. Moreover, sense 4 had 8 instances in the current BNC data:

(39) **EA5 1445** . . . one day he would **forget** how efficient his kitchen hand was and just do away with her.

As mentioned above, sense 3 is also associated with *wh*-complements and accordingly, there were five instances of sense 3 in the BNC data:

(40) **EDU 768** . . . many revolutions died because their leaders could not **forget** how good they were at destruction.

6.6. *Forget* + reflexive

The reflexive was found 8 times among the tokens drawn from the BNC. The frequency of the reflexive complement has actually been on the rise in the previous data (cf. table 16), but in the BNC the normed frequency count is again lower. In the first part of the CLMET there were no occurrences at all while in the later data the reflexive is found in CLMET parts 2 and 3 as well as in the BNC. The position seems now somewhat stable, although the frequency is admittedly rather low.

All of the tokens with reflexive complements were idiomatic and conformed to sense 5 'to forget oneself' and more precisely 5b 'to lose remembrance of one's own station, position or character':

(41) **JY7 4424** 'You **forget** yourself,' he said in a low, threatening voice.'You are in my home, *cara*. I make the rules here, no one else. You will leave when I say you may, not a moment sooner!'

(42) **CKC1006** Deplorable, deplorable. I **forget** myself. Such manners. Should have been ladies first, of course.

There were no cases of any extra-semantic constraints, insertions or extractions. It seems that the complements of *forget* involving reflexive pronouns are indeed very stable in character and consistently only appear in sense 5. However, it must be noted here that also other complements can appear in sense 5, especially NPs, even though the sense is more literal in those cases and there is a lack of idiomaticity. Thus the relationship between sense and complement is not as dogmatic as the relationship between complement and sense. All the reflexive pronoun complements found among

the corpora data were in fact representatives of sense 5, and as there were no exceptions, this seems to be a firm rule rather than a tendency.

6.7 *Forget* + gerund

There were only two tokens with the pattern *forget* + *-ing* -form in the BNC sample:

(43) **EDJ 1269** Forget the flourish of the big white napkin, forget two thumbs on the cork's corona, **forget** aiming at the bulbs in the recessed ceiling-lights. Just hold the cork and turn the bottle.

(44) **H8A 34** Never would he **forget** sitting beside her in the night air watching an open-air performance of The Parvenue and listening to the band at Fort House on Thursday evening.

In addition to the two self-evident cases of the gerund complement, there were some special cases as well. Quirk et al (1985, 1193) claim that the *-ing*-clause is rare without preceding *about* and consequently, the pattern *forget* + *about* + *-ing*-clause is in fact found among the data as well even though there were no evidence of it in the CLMET samples. However, there were still only two instances to be found, and they were discussed in more detail in subsection 6.3 and thus not included here. There was also one instance of the pattern *forget* + NP + V-*ing*:

(45) **CAK 1292** Who can **forget** the Catholic homosexual dipsomaniac being hit on the head by a dead seagull in *The Fourth Man* ?

As for the senses of the tokens, example (43) represents sense 4, example (44) sense 1 and example sentence (45) sense 1. The tokens found among the CLMET data all had sense 1, too, and there was an auxiliary present. Furthermore, there is also an auxiliary verb present in all of the other instances than example sentence (43), which suggests that if an auxiliary precedes the higher verb *forget* in cases with a gerundial complement, the appropriate sense is sense 1. All three tokens

were in accordance with the horror aequi principle and there were no other extra-semantic constraints, either.

The amount of gerunds conforms to the earlier findings, since there were no instances to be found among the data from CLMET part 1 and 3 and there were only a few examples in CLMET part 2 as well as in the present BNC data. Consequently, it seems that *forget* very rarely selects gerundial complements and the infinitival form is preferred instead, which is in conflict with the Great Complement Shift. As there were no extractions or insertion present either with gerundial or with the infinitival complements, it remains unclear how they would affect the selection between the two complements. However, the speculation seems somewhat redundant in the case of *forget*, since the data from the corpora suggest that structural discontinuity is rare with *forget* in any case.

It is curious that there are so few gerunds in the corpus data, since all the grammars and dictionaries state that *forget* can take *-ing*-clause complements. The frequency has remained low throughout the period under investigation in this thesis, but in the most recent data, i.e. data drawn from the BNC, there is some evidence of other gerundial patterns than just the simple *forget* + V-*ing* pattern, as can be detected from example (44). Although the frequencies are low and there are only a few instances available in present data, it still implies that there is a change in progress. It would be interesting to study the issue further, but unfortunately it exceeds the scope of this thesis as the BNC is chosen to represent the most recent data in the present thesis and thus the diachronic aspect does not reach any further.

6.8 Summary of the corpora results

After analyzing and discussing data from all parts of the CLMET as well as from the BNC, a brief recap is in order. Figure 1 shows the development in the frequencies of

the different complementation patterns over the period under investigation. The columns represent the normalized frequency counts in all of the corpora used as a source of data for the present thesis:

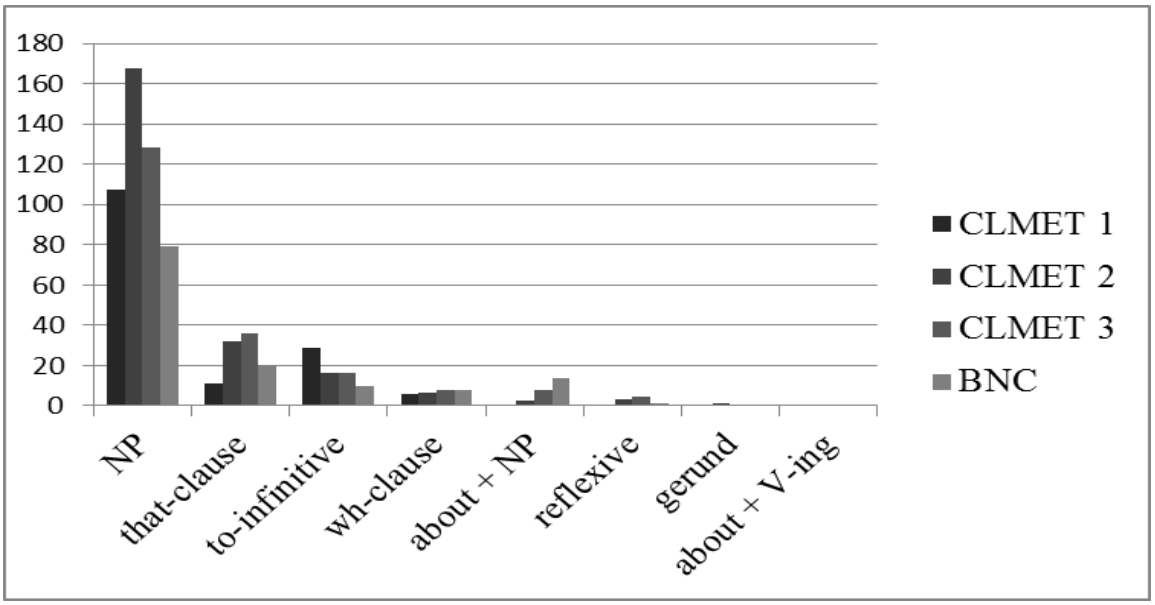

Figure 1. Normed frequency counts of different complements in all corpora.

It is evident that the NP is the most common complement in all the corpora and the difference in figures is truly extensive. Another non-sentential complement that seems to be thriving, is the pattern *about* + NP which seems to be gaining ground. However, the distribution of the other complements is more subtle in places. For instance, in the earliest data *to*-infinitives were far more common than *that*-clauses whereas some of the complements, such as the gerund, are hardly, if at all, discernible in the earlier data. The sentential complement patterns seem to be undergoing a change, as the frequency of infinitival complements is declining readily, the proportion of *that*-clauses seems to be unstable and in constant change, *wh*-clauses are on the rise and the gerund makes an appearance every now and then. At times it seemed that *forget* follows the principles of the Great Complement Shift, as the normed frequency count of infinitival complements decreased and the amount of gerundial complements rose in the second part of the CLMET. When looking from a wider perspective though, the

theory is after all not very well in accordance with the findings, as according to it gerunds should replace infinitives. The infinitives are admittedly on decrease, but they are not replaced by gerunds but rather with other sentential compelements.

Considering what the secondary sources suggested about the complementation of *forget*, there are both consistencies and deviations discernible. Firstly, only the dictionaries mention the NP complement, which is the most common complement in the authentic data. *To*-infinitives are offered by all of the secondary sources, and yet in authentic usage the proportion is declining notably.

As to the senses that *forget* takes, the following figure displays the development of the different senses in chronological order:

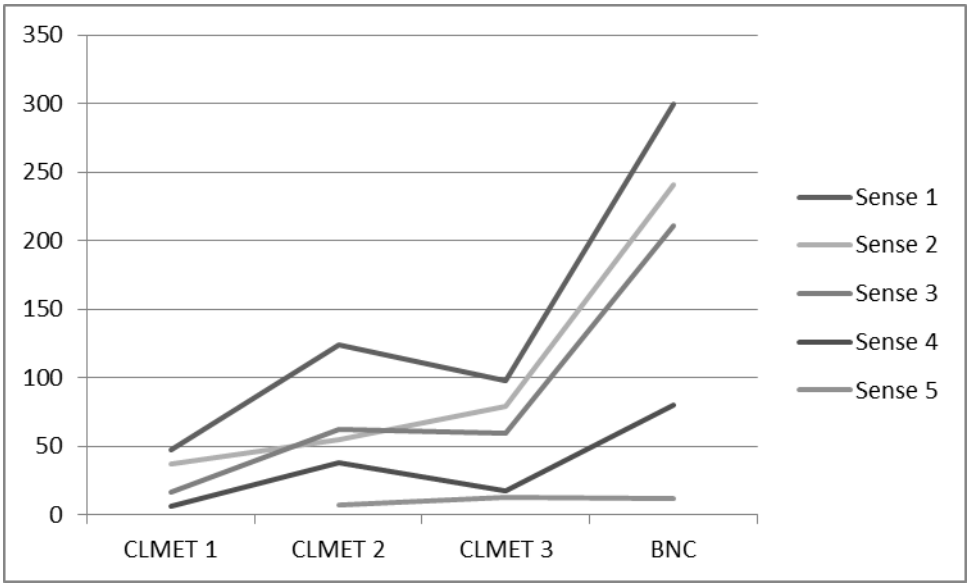

Figure 2. Distribution of different senses in the corpora.

Sense 1'not to remember or fail to recollect' is advancing, whereas sense 5 'to forget oneself' is becoming even more rare. Senses 2 'to omit or neglect through inadvertence' and 3 'to cease or omit to think of' are both increasing in amount, the frequency counts being close to each other and the mutual order changing over the course of time. Sense 4 'to disregard or overlook' has an established role, but is still somewhat rare compared to the senses 1, 2 and 3.

7 Conclusion

The aim was to study the complements of the verb *forget* in Late Modern English with the help of corpus data from the CLMET and the BNC. The most common complement was NP, although it was not mentioned by any grammarians. The proportion of the NP complement was notable in both parts of the corpus. Although the secondary sources offered a total of 7 possible complements for *forget*, only four different complements were found in the CLMET part 1 data, whereas all seven are present in the data from the second part of the corpus. Since the data from the first part of the CLMET is from 1710-1780 and the secondary sources from 20th and 21st centuries, this might suggest that the use and complementation patterns of the verb *forget* have developed so that there is now more variation in the complements *forget* selects than in the 18th century. The fact that Poutsma's *Grammar of Late Modern English* (1904) only mentions one complement supports the supposition that the variety of complements has been increasing over time, and so does also the fact that the later parts of the CLMET (texts from 1780-1850 and 1850-1920) offers more alternatives than the first part (1710-1780), not to speak of the BNC, which introduces even more complements, such as the pattern *about* + V-*ing* which is completely absent from the CLMET material.

To a certain extent the theory of the Great Complement Shift is possibly applicable, as the infinitival complement seemed to be on the decrease and the gerund variant seemed to be slowly emerging based on the first two parts of the CLMET. However, as the study proceeded towards the modern day usage of *forget*, the development ceased to be as clear. It is evident that infinitives are decreasing, but the gerund is not advancing as it should according to the theory. Instead, *that*-clauses have increased, excluding the BNC data, and the proportion of *wh*-clauses is also slowly growing. Some further research would be needed to examine the spread of the

gerund or other clausal complements on the expense of the infinitive as regards *forget*, so as to investigate whether the predictions of the Great Complement Shift halted altogether in the late 18th century or whether the change is still in progress, albeit it being slow and obscure. Even if the changes are not always radical and straightforward in nature, the Great Complement Shift has clearly affected the complementation system of *forget* in recent centuries.

Furthermore, according to the OED only sense 2 is appropriate for infinitival complements, which in fact is in perfect accordance with the corpus data. Interestingly almost all gerund complements also have sense 2, which violates against Bolinger's principle of different grammatical structures having a different meaning. However, consistent with the earlier findings, towards modern days more variation is detected in the use of gerund complements. Furthermore, verbs of communication were particularly common in sentences with *to*-infinitive complements. The tendency to allow for more senses applies to other complements as well, such as the pattern *about* + NP, which was always in sense 3 in the CLMET data whereas in the BNC also sense 1 is found. According to Rohdenburg's Complexity Principle (1996, 151) "in the case of more or less explicit grammatical options the more explicit one(s) will tend to be favored in cognitively more complex environments". Consequently, in the earlier data there was a clear tendency to retain that in *that*-clauses with insertion, while insertion adds to the complexity of the sentence and retaining *that* is the more explicit grammatical option as opposed to the less explicit choice of omitting *that*. However, the principle begins to lose its charm towards the modern days, and consequently there are relatively many cases of insertion in sentences where *that* is omitted in the BNC data. It has to be noted here, that the results are indicative rather than inclusive, since the results of a corpus study "can only be generalized to the extent that the corpus is a representative sample" (Ball 1994, 295, cf. 2.2). However,

as Mair states in section 2.1, it is still important to study the changes in frequency counts of the variants and document the gradual spread of the constructions (2002, 108).

As regards the extra-semantic constraints and their effect on the data, there were some constraints clearly noticeable, apart from the horror aequi principle which proved to be almost completely absent from the empirical data. However, the horror aequi principle was still applicable in that there were hardly any violations to it as there were only three tokens with two adjacent *to*-infinitives. The violations appeared in the third part of the CLMET and in the BNC which suggests that the principle is in more recent days not as effective as it used to be, although the frequencies of such violations are still so low that it is hard to tell whether it is still a mere coincidence or is there in fact a change in progress. Moreover, extractions are especially common among the NP complements, and the frequency of extractions is increasing. Intriguingly extractions appear most frequently with sense 1 NP complements, whereas insertions tend to be attracted by sense 3. There were hardly any insertions with NP complements in sense 1 and hardly any extractions among tokens relating to sense 3. Consequently, there are clear tendencies to be noted in the relationships between senses and complements, although some of them are more transparent while others are more obscure.

Bibliography

Primary Source:

The British National Corpus

The Corpus of Late Modern English Texts

Secondary Sources:

Ball, C. N. 1994. "Automated Text Analysis: Cautionary Tales." *Literary and Linguistic Computing* 9, 295-302.

Biber, D. et al. 1998. *Corpus Linguistics. Investigating language structure and use*. Cambridge: Cambridge University Press.

Biber, D. et al. 1999. *Longman Grammar of Spoken and Written English*. Harlow: Longman.

Bolinger, D. 1968. "*Entailment and the meaning of structures.*" *Glossa* 2: 119-127.

Collins Cobuild English Language Dictionary. 1987. Ed. J. Sinclair. London: Collins.

De Smet, H. 2005. "A Corpus of Late Modern English Texts." *ICAME Journal* 29: 69-82.

Herbst, T. et al 2004. *A Valency Dictionary of English*. Berlin: de Gruyter.

Huang, C.-T. J. 1997. "Introduction to Syntax." LSA Linguistic Institute.

Huddleston, R. 1984. *Introduction to the Grammar of English*. Cambridge: Cambridge University Press.

Huddleston, R. and Pullum, G. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Leech, G. 1968. "Some assumptions in the metatheory of linguistics". *Linguistics* 39: 87-102.

Leech, G. and Svartvik, J. 2002. 3rd ed. *A Communicative Grammar of English*. Harlow: Pearson Education Limited.

Mair, C. 2002. "Three changing patterns of verb complementation in Late Modern English: a realtime study based on matching text corpora." *English Language and Linguistics* 6, 1: 105-131.

http://www.natcorp.ox.ac.uk/docs/URG.xml

The Oxford English Dictionary. 1989. 2nd ed. Oxford: Oxford University Press. http://dictionary.oed.com/ [Accessed March 31st 2008]

Poutsma, H. 1904. *A Grammar of Late Modern English*. Groningen: Noordhoof.

Quirk, R. et al. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.

Rohdenburg, G. 1996. "Cognitive complexity and increased grammatical explicitness in English." *Cognitive Linguistics* 7, 2: 149-182.

Rohdenburg, G. 2003. "Cognitive complexity and *horror aequi* as factors determining the use of interrogative clause linkers in English." In *Determinants of grammatical variation in English*. Eds. Rohdenburg, Mondorf. Berlin: Mouton de Gruyter.

Rohdenburg, G. 2006. "The Role of Functional Constraints in the Evolution of the English Complementation System." In *Syntax, Style and Grammar Norms.* Eds. Dalton-Puffer, Kastovsky, Ritt and Schendl. Bern: Peter Lang.

Somers, H. L. 1984. "On the validity of the complement-adjunct distinction in valency grammar." *Linguistics* 22, 22: 507-530.

Somers, H.L. 1987. *Valency and Case in Computational Linguistics*. Edinburgh: Edinburgh University Press.

Svartvik, J. 1992. "Corpus linguistics comes of age." In *Trends in Linguistics*. Ed. Winter. Berlin, New York: Mouton de Gruyter.

Vosberg, U. 2003a. "The role of extractions and horror aequi in the evolution of –ing complements in Modern English". In *Determinants of Grammatical Variation in English*. Eds. Rohdenburg, Mondorf. Berlin, New York: Mouton de Gruyter.

Vosberg, U. 2003b. "Cognitive Complexity and the Establishment of –ing Constructions with Retrospective Verbs in Modern English." In *Insights into Late Modern English.* Eds. Dossena and Jones. Bern: Peter Lang.

Vosberg, U. 2009. "One Language, Two Grammars?" In *Studies in English Language.* Eds. Rohdenburg and Schlüter. Cambridge: Cambridge University Press.